

# **Data Compression for Helioseismology**

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

“Doctor rerum naturalium”

der Georg-August-Universität Göttingen

im Promotionsprogramm PROPHYS

der Georg-August University School of Science (GAUSS)

vorgelegt von

**Björn Löptien**

aus Brunsbüttel

Göttingen, 2015

### Betreuungsausschuss

Dr. Aaron C. Birch

Max-Planck-Institut für Sonnensystemforschung, Göttingen

Prof. Dr. Laurent Gizon

Max-Planck-Institut für Sonnensystemforschung, Göttingen

Institut für Astrophysik, Georg-August-Universität Göttingen

Prof. Dr. Sami K. Solanki

Max-Planck-Institut für Sonnensystemforschung, Göttingen

### Mitglieder der Prüfungskommission

Referent: Prof. Dr. Laurent Gizon

Max-Planck-Institut für Sonnensystemforschung, Göttingen

Institut für Astrophysik, Georg-August-Universität Göttingen

Korreferent: Prof. Dr. Manfred Schüssler

Max-Planck-Institut für Sonnensystemforschung, Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Andreas Tilgner

Institut für Geophysik, Georg-August-Universität Göttingen

Prof. Dr. Ansgar Reiners

Institut für Astrophysik, Georg-August-Universität Göttingen

Prof. Dr. Jörg Büchner

Max-Planck-Institut für Sonnensystemforschung, Göttingen

Prof. Dr. Stefan Dreizler

Institut für Astrophysik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 29. Juli 2015

---





# Summary

Efficient data compression will play an important role for several upcoming and planned space missions involving helioseismology, such as *Solar Orbiter*. *Solar Orbiter*, to be launched in October 2018, will be the next space mission involving helioseismology. The main characteristic of *Solar Orbiter* lies in its orbit. The spacecraft will have an inclined solar orbit, reaching a solar latitude of up to  $33^\circ$ . This will allow, for the first time, probing the solar poles using local helioseismology. In addition, combined observations of *Solar Orbiter* and another helioseismic instrument will be used to study the deep interior of the Sun using stereoscopic helioseismology. The Doppler velocity and continuum intensity images of the Sun required for helioseismology will be provided by the *Polarimetric and Helioseismic Imager* (PHI).

Major constraints for helioseismology with *Solar Orbiter* are the low telemetry and the (probably) short observing time. In addition, helioseismology of the solar poles requires observations close to the solar limb, even from the inclined orbit of *Solar Orbiter*. This gives rise to systematic errors.

In this thesis, I derived a first estimate of the impact of lossy data compression on helioseismology. I put special emphasis on the *Solar Orbiter* mission, but my results are applicable to other planned missions as well.

First, I studied the performance of PHI for helioseismology. Based on simulations of solar surface convection and a model of the PHI instrument, I generated a six-hour time-series of synthetic Doppler velocity images with the same properties as expected for PHI. Here, I focused on the impact of the point spread function, the spacecraft jitter, and of the photon noise level. The derived power spectra of solar oscillations suggest that PHI will be suitable for helioseismology.

The low telemetry of *Solar Orbiter* requires extensive compression of the helioseismic data obtained by PHI. I evaluated the influence of data compression using data provided by the *Helioseismic and Magnetic Imager* (HMI). HMI is an instrument onboard the *Solar Dynamics Observatory* (SDO), launched in 2010. It provides full disk images with high cadence of the continuum intensity, the Doppler-velocity and the full magnetic field vector. Using HMI Doppler-velocity maps, I showed that the signal-to-noise ratio of supergranulation in time-distance helioseismology is robust regarding lossy data compression. In addition, I proved that the accuracy and precision of probing differential rotation with local correlation tracking of granulation is not severely affected by extensive lossy data compression. This indicates that the low telemetry of *Solar Orbiter* may not be a major challenge for helioseismology.

# Zusammenfassung

Die effiziente Kompression von Daten wird eine wichtige Rolle für mehrere bevorstehende und geplante Weltraummissionen spielen, die Helioseismologie betreiben werden, wie beispielsweise *Solar Orbiter*. *Solar Orbiter* ist die nächste Mission, die Helioseismologie beinhaltet, und soll im Oktober 2018 gestartet werden. Das Hauptmerkmal von *Solar Orbiter* ist der Orbit. Die Umlaufbahn des Satelliten wird zur Ekliptik geneigt sein, sodass der Satellit einen solaren Breitengrad von bis zu  $33^\circ$  erreichen wird. Dies wird erstmals ermöglichen, die Pole der Sonne mit Hilfe von lokaler Helioseismologie zu studieren. Zusätzlich dazu können kombinierte Beobachtungen von *Solar Orbiter* und einem anderen Instrument dazu benutzt werden, die tiefen Schichten der Sonne mittels stereoskopischer Helioseismologie zu erforschen. Die Aufnahmen der Dopplergeschwindigkeit und der Kontinuumsintensität, die für Helioseismologie benötigt werden, werden vom *Polarimetric and Helioseismic Imager* (PHI) geliefert werden.

Große Hindernisse für Helioseismologie mit *Solar Orbiter* sind die niedrige Datenübertragungsrate und die (wahrscheinlich) kurzen Beobachtungszeiten. Außerdem erfordert die Untersuchung der Pole der Sonne Beobachtungen in der Nähe des Sonnenrandes, sogar von dem geneigten Orbit von *Solar Orbiter* aus. Dies kann zu systematischen Fehlern führen.

In dieser Doktorarbeit gebe ich eine erste Einschätzung ab, wie stark Helioseismologie von verlustbehafteter Datenkompression beeinflusst wird. Mein Schwerpunkt liegt dabei auf der *Solar Orbiter* Mission, die von mir erzielten Ergebnisse sind aber auch auf andere geplante Missionen übertragbar.

Zunächst habe ich mit Hilfe synthetischer Daten die Eignung des PHI Instruments für Helioseismologie getestet. Diese basieren auf Simulationen der Konvektion nahe der Sonnenoberfläche und einem Modell von PHI. Ich habe eine sechs Stunden lange Zeitreihe synthetischer Daten erstellt, die die gleichen Eigenschaften wie die von PHI erwarteten Daten haben. Hierbei habe ich mich auf den Einfluss der Punktspreizfunktion, der Vibrationen des Satelliten und des Photonenrauschen konzentriert. Die von diesen Daten abgeleitete spektrale Leistungsdichte der solaren Oszillationen legt nahe, dass PHI für Helioseismologie geeignet sein wird.

Aufgrund der niedrigen Datenübertragungsrate von *Solar Orbiter* müssen die von PHI für die Helioseismologie gewonnenen Daten stark komprimiert werden. Ich habe den Einfluss von Kompression mit Hilfe von Daten getestet, die vom *Helioseismic and Magnetic Imager* (HMI) stammen. HMI ist ein Instrument an Bord des *Solar Dynamics Observatory* Satelliten (SDO), der 2010 gestartet worden ist. HMI erstellt mit hoher zeitlicher Abfolge Karten der Kontinuumsintensität, der Dopplergeschwindigkeit und des kompletten Magnetfeldvektors für die komplette von der Erde aus sichtbare Hemispäre der Sonne. Mit Hilfe mit von HMI aufgenommenen Karten der Dopplergeschwindigkeit

---

konnte ich zeigen, dass das Signal-zu-Rausch Verhältnis von Supergranulation in der Zeit-Entfernungs Helioseismologie nicht stark von Datenkompression beeinflusst wird. Außerdem habe ich nachgewiesen, dass die Genauigkeit und Präzision von Messungen der Sonnenrotation mittels Local Correlation Tracking von Granulation durch verlust-behaftete Datenkompression nicht wesentlich verschlechtert werden. Diese Ergebnisse deuten an, dass die niedrige Datenübertragungsrate von *Solar Orbiter* nicht unbedingt ein großes Hinderniss für Helioseismologie sein muss.



# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Motivation . . . . .	11
1.2	Outline . . . . .	12
1.3	Introduction to Helioseismology . . . . .	12
1.3.1	Solar Oscillations . . . . .	12
1.3.2	Time-Distance Helioseismology . . . . .	13
1.4	The Solar Orbiter Mission . . . . .	16
1.4.1	Mission Design . . . . .	16
1.4.2	Orbit . . . . .	16
1.4.3	Instruments . . . . .	18
1.4.4	The Polarimetric and Helioseismic Imager (PHI) . . . . .	18
1.4.5	Science Goals Involving Helioseismology . . . . .	23
1.4.6	Observing Strategies for Helioseismology . . . . .	25
1.4.7	Data Compression . . . . .	27
<b>2</b>	<b>Simulating PHI Time-Series for Helioseismology</b>	<b>33</b>
2.1	Introduction to Chapter 2 . . . . .	33
2.2	Steps in the Generation of Synthetic Data . . . . .	33
2.2.1	Simulations of Solar Convection . . . . .	33
2.2.2	Computation of Line Profiles . . . . .	34
2.2.3	Simulation Tool: SOPHISM . . . . .	34
2.2.4	Using SOPHISM . . . . .	37
2.2.5	Spacecraft Jitter . . . . .	38
2.2.6	Point Spread Function . . . . .	40
2.2.7	Photon Noise . . . . .	40
2.3	Synthetic Intensity and Velocity Maps . . . . .	42
2.4	Oscillation Power Spectra . . . . .	42
<b>3</b>	<b>Image Compression in Local Helioseismology</b>	<b>47</b>
3.1	Introduction to Chapter 3 . . . . .	47
3.2	Compression Schemes . . . . .	48
3.2.1	Quantization and Huffman Encoding . . . . .	48
3.2.2	JPEG Compression . . . . .	49
3.2.3	Smoothing and Subsampling . . . . .	49
3.3	Results . . . . .	50
3.3.1	Input Observations . . . . .	50

3.3.2	Influence on Helioseismic Power Spectra . . . . .	51
3.3.3	Influence on Supergranulation Travel Times . . . . .	54
3.4	Discussion and Conclusion for Chapter 3 . . . . .	57
<b>4</b>	<b>Data Compression for Local Correlation Tracking of Solar Granulation</b>	<b>61</b>
4.1	Introduction to Chapter 4 . . . . .	61
4.2	Data and Methods . . . . .	62
4.2.1	Input Data . . . . .	62
4.2.2	The FLCT Code . . . . .	62
4.2.3	Compression Methods . . . . .	63
4.2.4	Calibration of LCT Velocities . . . . .	64
4.3	Results . . . . .	64
4.3.1	Individual Flow Maps . . . . .	64
4.3.2	Shrinking-Sun Effect . . . . .	65
4.3.3	Differential Rotation . . . . .	71
4.3.4	HMI $P$ -Angle Oscillations . . . . .	71
4.3.5	Influence of Compression on Differential Rotation . . . . .	72
4.4	Discussion and Conclusion for Chapter 4 . . . . .	75
<b>5</b>	<b>Discussion and Outlook</b>	<b>79</b>
5.1	Feasibility of Helioseismic Science Goals . . . . .	79
5.2	Outlook: Probing Differential Rotation with Time-Distance Helioseis- mology . . . . .	80
5.3	Outlook: Modeling the Shrinking-Sun Effect . . . . .	81
5.4	Future Work . . . . .	83
	<b>Bibliography</b>	<b>85</b>
	<b>Publications</b>	<b>95</b>
	<b>Acknowledgements</b>	<b>97</b>

# 1 Introduction

## 1.1 Motivation

Several upcoming and proposed space missions involving helioseismology will suffer from low telemetry and thus, require extensive data compression. The next helioseismology mission is *Solar Orbiter*, which is scheduled for launch in October 2018. It was selected in 2011 as an ESA M-class mission and is the successor of the highly successful SOHO mission of ESA and NASA. *Solar Orbiter* will be the first mission to perform helioseismology outside of the ecliptic. The spacecraft will have an elliptic and inclined orbit with a minimum distance to the Sun of 0.28 AU and a maximum solar latitude of  $33^\circ$ . This allows observing the solar poles from a much lower angle than currently possible.

The main scientific goal of the mission is to study the connection between the Sun and the heliosphere. This requires a combined analysis including results obtained from different instruments, both in-situ and remote-sensing, and the application of different methods. A key science objective is probing flows in the solar interior with helioseismology using time-series of the Doppler-velocity provided by the *Polarimetric and Helioseismic Imager* (PHI). Mapping large scale and convective flows at high latitudes using helioseismology will offer important ingredients for modeling the solar dynamo. In addition, *Solar Orbiter* will test the concept of stereoscopic helioseismology - observing solar oscillations from two different vantage points using two instruments. This potentially will allow studying the deep interior of the Sun, including the tachocline.

Helioseismology with *Solar Orbiter* will be subject to challenges that are not relevant for current helioseismology missions. The observing time for helioseismology will probably be limited. This puts constraints on the noise level that can be achieved. In addition, helioseismology will suffer from the low telemetry of the spacecraft, requiring extensive data compression. Further challenges might arise from systematic errors. Many science objectives of *Solar Orbiter* involve observations at high solar latitudes. Although the orbit of the spacecraft will be inclined to the ecliptic, observing the poles of the Sun might be affected by various systematic effects in helioseismology that occur when observing far from disk center. Another concern is the varying distance of the spacecraft from the Sun. This could change the thermal environment of the PHI instrument and influence the performance of the instrument, e.g., the point spread function.

These issues demand the development of detailed observing strategies for the individual science goals of *Solar Orbiter* involving helioseismology. The results of this thesis are a first step in this direction. I test the performance of the PHI instrument in the observation of solar oscillations and I evaluate the influence of data compression on selected measurements with time-distance helioseismology (Duvall et al. 1993) and local correla-

tion tracking of granulation (LCT, November & Simon 1988). Although I focus in this thesis on the *Solar Orbiter* mission, the results regarding compression are also applicable to other planned missions, such as L5 or SAFARI (for a review of concepts for future missions see Sekii et al. 2015).

## 1.2 Outline

The introduction is divided into two parts. The first part gives a short summary on solar oscillations and how they are used in helioseismology to probe the interior of the Sun. The second part of the introduction presents the *Solar Orbiter* mission. It summarizes the opportunities offered by the mission for helioseismology and the challenges that are connected to this. In the second chapter, I discuss how the observation of solar oscillations is affected by the PHI instrument, e.g., by the point spread function. I generated a time-series of synthetic data with the same properties as expected by the High Resolution Telescope of PHI and evaluated the impact of different instrumental effects on the power spectrum of solar oscillations. In the third and fourth chapters, I test the influence of data compression on measuring flows with *Solar Orbiter*. In the third chapter, I applied different compression methods to tracked and remapped Dopplergrams provided by the *Helioseismic and Magnetic Imager* (HMI, Schou et al. 2012) and evaluated the influence of compression on helioseismic power spectra and the signal-to-noise ratio of supergranulation in travel time maps. The fourth chapter discusses the influence of compression on the accuracy and precision of determining differential rotation using LCT. The discussion summarizes the results presented in the individual chapters and gives an outlook about future work that needs to be done in order to prepare for helioseismology with *Solar Orbiter*.

## 1.3 Introduction to Helioseismology

### 1.3.1 Solar Oscillations

Time-series of observations of the Sun in e.g., Doppler-velocity or continuum intensity, reveal that the Sun exhibits oscillations with a period of about five minutes. These oscillations were first discovered by Leighton et al. (1962) and with the help of additional observations (Deubner 1975, Claverie et al. 1979) and theoretical work (Ulrich 1970, Leibacher & Stein 1971), they were interpreted as being the superposition of a large number of normal modes that are stochastically excited by turbulent convection. These normal modes can be divided in two groups. The majority of the oscillations are acoustic modes (called p-modes) that penetrate into the interior of the Sun. In addition, there are surface gravity waves (called f-modes) at the solar surface, where the restoring force is buoyancy (similar to deep ocean waves). There have also been claims about the existence of gravity waves (g-modes) in the radiative zone of the Sun (e.g., García et al. 2007), but their existence is still an issue of debate. If they exist, they are evanescent in the solar convection zone, and so, they would have an extremely low amplitude in observations of the Doppler-velocity or continuum intensity of the photosphere.



The horizontal component of the eigenmodes is given by spherical harmonics with degree  $\ell$  and azimuthal order  $m$ . F-modes are confined close to the solar surface and decay exponentially with depth. P-modes, on the other hand, also exhibit a series of nodes in radial direction. The number of nodes defines the radial order  $n$  of the normal mode. P-modes penetrate to different depths in the Sun, depending on the radial order  $n$  and the degree  $\ell$ .

Since the solar oscillations are stochastically excited, the resulting wave field is a superposition of many individual normal modes. This can be seen, when plotting the power of the observed wave field as a function of the degree  $\ell$  and the frequency  $\nu$ . Such a plot is called power spectrum (see Figure 1.1 for an example). There are two distinct features in this plot. At low frequencies, there is an excess in power originating from convection (granulation). At frequencies around 3 mHz, there are several ridges, each of them consisting of hundreds of individual normal modes. The ridge at lowest frequencies consists of f-modes, the other ones are p-modes with increasing values of the radial order  $n$  (named  $p_1$ ,  $p_2$ , etc.). See e.g., Christensen-Dalsgaard (2003) for more information on solar oscillations.

In helioseismology, these modes are used to probe the interior of the Sun, similar to seismology of the Earth. Global helioseismology is based on accurate measurements of the mode frequencies, which depend on the sound speed in the interior of the Sun. In addition, deviations from spherical symmetry, as introduced for example by differential rotation, lift the  $\ell$ -degeneracy of the normal modes. Since the individual normal modes have different dependencies on depth and latitude, a combined analysis of modes with different values of  $\ell$ ,  $m$ , and  $n$  allows us to infer properties of the Sun as a function of radius and latitude. Global helioseismology has led to important discoveries about the interior structure of the Sun. It has been used to measure the sound speed as a function of solar radius (e.g., Christensen-Dalsgaard et al. 1985), which allowed determination of the depth of the convection zone and also helped solve the solar neutrino problem. In addition, global helioseismology has mapped solar rotation throughout most of the solar interior (e.g., Schou et al. 1998, Thompson et al. 2003) and has determined the helium abundance in the convection zone (e.g., Kosovichev et al. 1992).

Local helioseismology makes use of the full wave field observed at the solar surface. This allows us to obtain 3D information about the solar interior. Local helioseismology can be used to study a large variety of solar phenomena. This includes active regions, convection, meridional circulation and solar rotation. Several different methods are used in local helioseismology, one of them is time-distance helioseismology.

### 1.3.2 Time-Distance Helioseismology

Time-distance helioseismology (Duvall et al. 1993) infers properties of the solar interior by measuring the time it takes waves to travel between two points on the solar surface. These wave travel times depend on the structure of the solar interior, e.g., on changes of the sound speed or flows. The travel times of waves propagating in the direction of a flow are decreased, for waves propagating in the opposite direction, the travel times are increased. The difference between these two travel times allows inferring the properties of the underlying flow. Which flows are being probed depends on the geometry of the measurement. One option is to determine the travel time of waves propagating between a

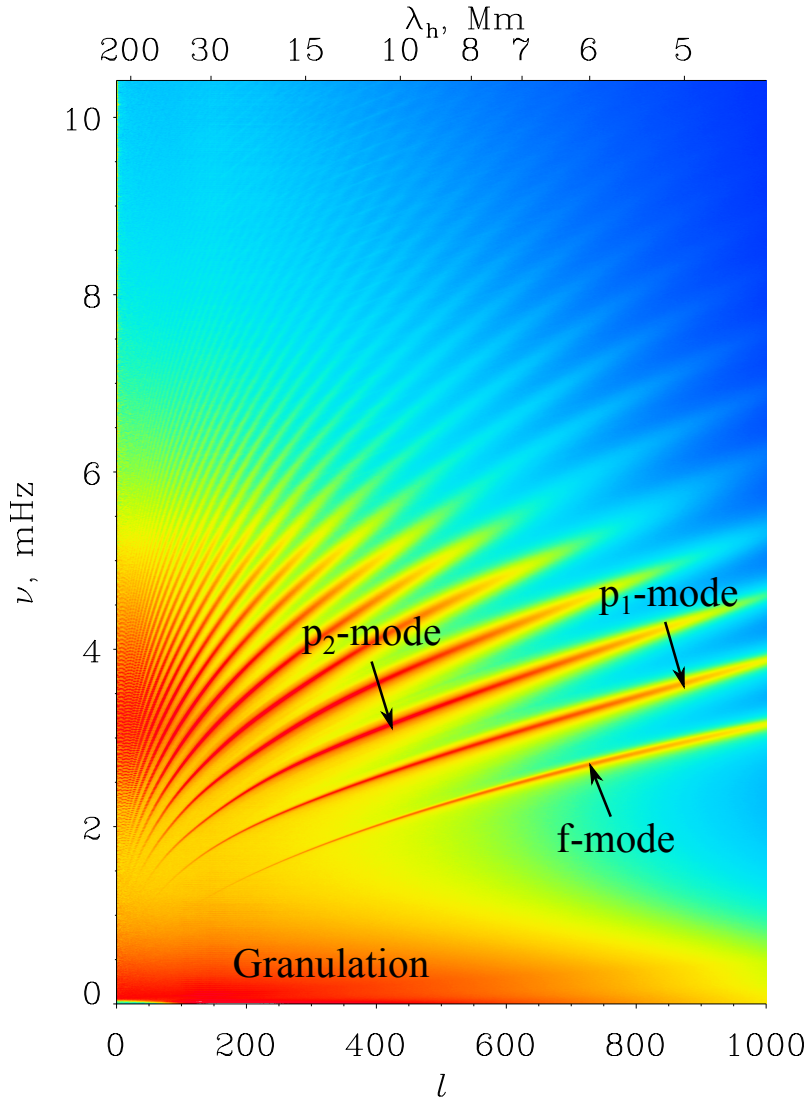


Figure 1.1: Power Spectrum of solar oscillations derived from data provided by the *Michelson Doppler Imager* (MDI, Scherrer et al. 1995). Courtesy of the Stanford Solar Observatories group and the SOHO/MDI team. Reprinted with permission.

central point and a surrounding annulus. This allow measuring the horizontal divergence of the flows. Alternatively, the travel time between a central point and two quadrants east and west or north and south of it can be computed. Such a measurement is sensitive to flows in the east-west or north-south direction.

The travel times are derived from the temporal cross-covariance of the observed wave field (e.g., in Doppler velocity) on the solar surface. Figure 1.2 shows the amplitude of the cross-covariance of the solar oscillations as a function of the distance  $\Delta x$  and the time lag  $\Delta t$ . The cross-covariance function exhibits several ridges. The first ridge corresponds to waves propagating between the two points without any additional reflections at the solar surface, the other ridges correspond to waves that bounce of the surface while propagating between the two points. The larger the distance, the deeper the waves penetrate into the Sun. Since the sound speed increases with depth, the time lag does not increase linearly with  $\Delta x$ .

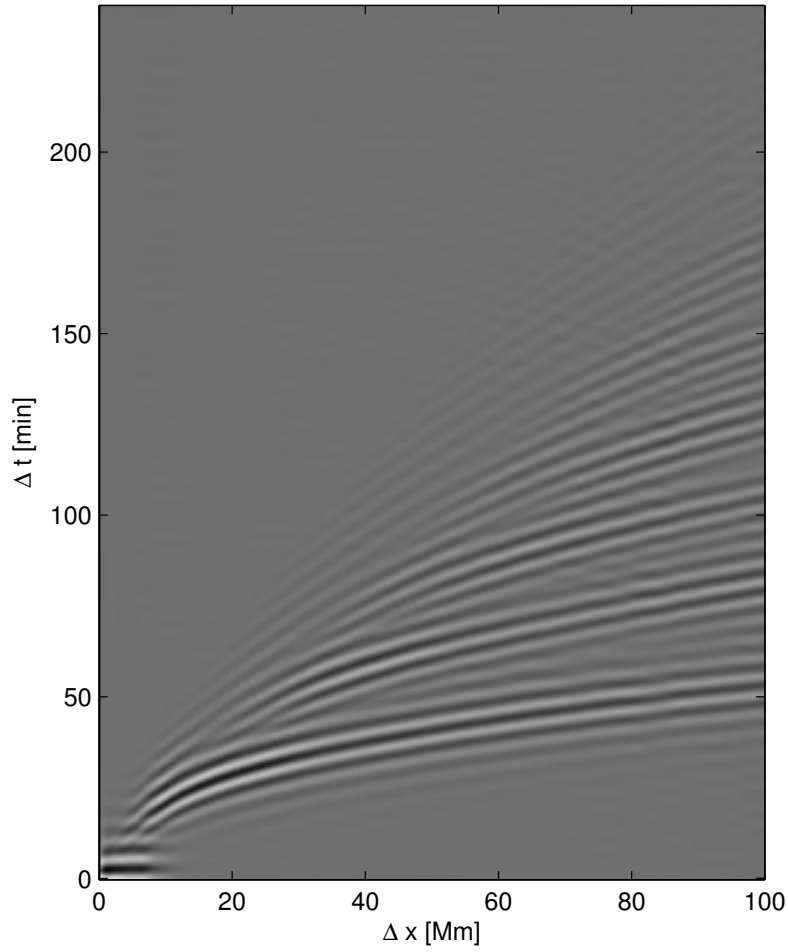


Figure 1.2: Observed cross-covariance of solar oscillations plotted as a function of time lag and distance. Such a plot is called time-distance diagram. The cross-covariances are derived from data provided by the *Helioseismic and Magnetic Imager* (HMI, Schou et al. 2012).

The cross-covariances shown in Figure 1.2 are hard to interpret, because they are affected by a large variety of modes that reach different depths in the Sun. The measurement can be focused on a given depth by computing the cross-covariances using a fixed distance  $\Delta x$ . Further constraints can be made by applying a filter to the wave field that selects only specific modes. In the last step, the travel times are determined from the cross-covariances. In this thesis, I use the method of Gizon & Birch (2002). It is a linear approach that compares the derived cross-covariances with a reference cross-covariance. The output of this method is neither the group nor the phase travel time, but the “travel times” are clearly defined. In case one is interested in the actual flows, an inversion of the travel times is required. See e.g., Gizon & Birch (2005) for more information on time-distance helioseismology and other methods used in local helioseismology.

## 1.4 The Solar Orbiter Mission

### 1.4.1 Mission Design

Scheduled for launch in October 2018, *Solar Orbiter* is the next space mission involving helioseismology. It was selected in 2011 as an ESA M-class mission and is the successor of the highly successful SOHO Mission of ESA and NASA. *Solar Orbiter* will be the first mission to perform helioseismology outside of the ecliptic. The spacecraft will have an elliptic and inclined orbit with a minimum distance to the Sun of 0.28 AU and a maximum solar latitude of  $33^\circ$ . This allows observing the solar poles from a much lower vantage point than currently possible.

The central science question of the mission is (Marsden & Müller 2011, Müller et al. 2013):

*How does the Sun create and control the heliosphere?*

This main objective can be divided into four top-level science goals:

- How and where do the solar wind plasma and magnetic field originate in the corona?
- How do solar transients drive heliospheric variability?
- How do solar eruptions produce the energetic particle radiation that fills the heliosphere?
- How does the solar dynamo work and drive the connections between the Sun and the heliosphere?

In order to answer these questions, *Solar Orbiter* will host both in-situ and remote-sensing instruments. Connecting in-situ measurements close to the Sun with remote-sensing observations of the different layers of the solar atmosphere is one of the key aspects of the mission. Each of the individual science objectives demands a combined analysis including results obtained from different instruments and the application of different methods. Especially the last science question requires knowledge about flows in the solar interior, which can be provided by helioseismology. Mapping large scale and convective flows at high latitudes using helioseismology will offer important ingredients for modeling the solar dynamo. In addition, *Solar Orbiter* will test the concept of stereoscopic helioseismology - observing solar oscillations from two different vantage points using two instruments. This potentially allows studying the deep interior of the Sun, including the tachocline. See also Löptien et al. (2014b) for a review of helioseismology with *Solar Orbiter*.

### 1.4.2 Orbit

*Solar Orbiter* will have a unique orbit (see Figure 1.3), allowing to study the Sun both from a close distance and from a vantage point that is inclined to the ecliptic. After launch, which is currently scheduled for October 2018, *Solar Orbiter* will undergo a 3-year cruise phase, during which no observations will be performed. During this time, the spacecraft will perform several Gravity-Assist-Maneuvers (GAMs) with Earth and Venus in order to lose angular momentum and decrease its distance to the Sun. After

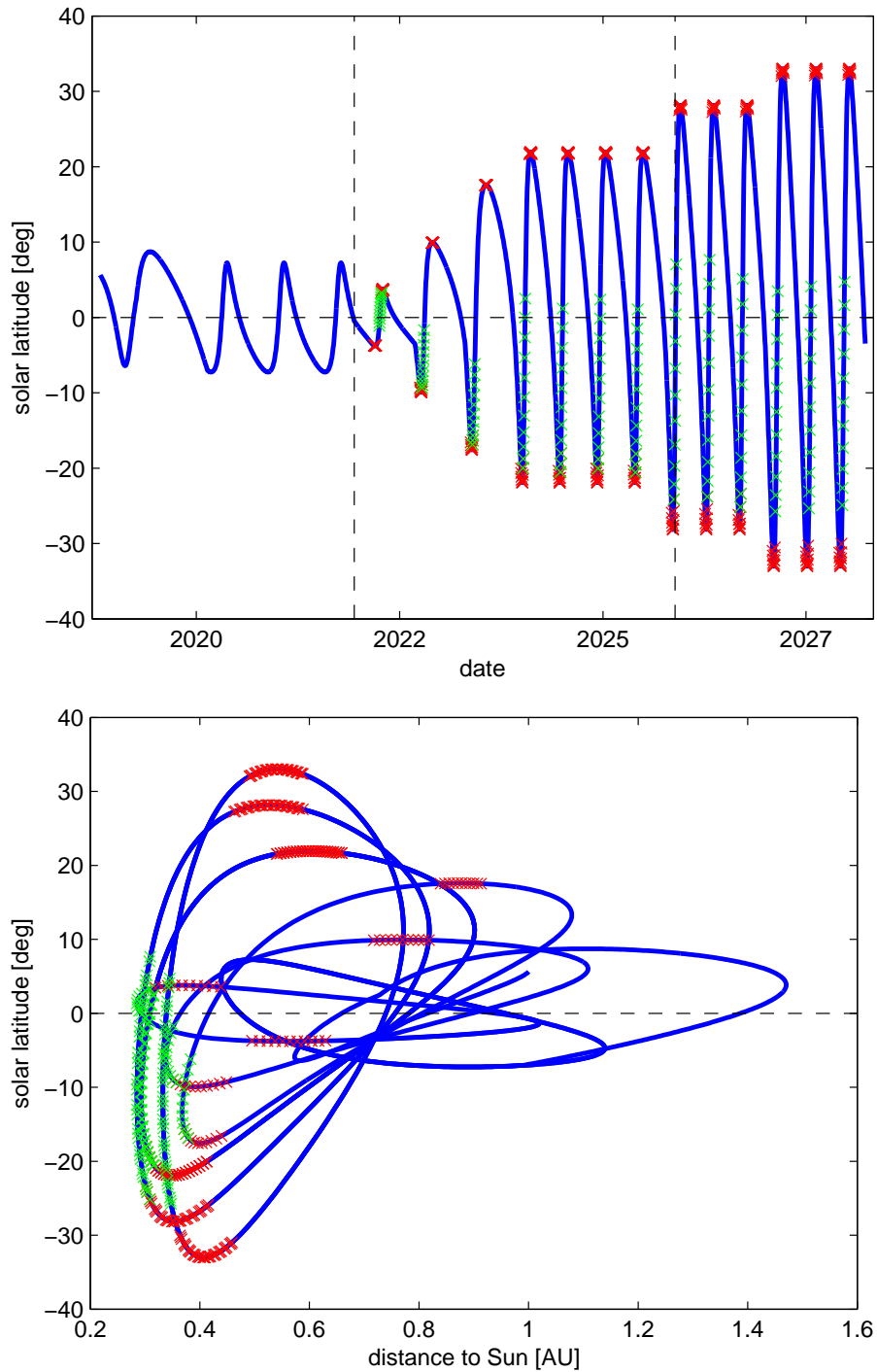


Figure 1.3: Orbit of *Solar Orbiter*. *Top*: Solar latitude of the spacecraft as a function of time, *bottom*: solar latitude as a function of the distance to the Sun. The vertical lines in the *top* plot show the beginning and the end of the nominal mission phase. *Solar Orbiter* will have a minimum distance to the Sun of 0.28 AU. The maximum inclination of the spacecraft will be  $28^\circ$  during the nominal mission phase and  $33^\circ$  during the extended mission. The current baseline for the remote-sensing instruments is to observe only during three science windows per orbit, at maximum northern/southern solar latitude (*red crosses*) and at perihelion (*green crosses*).

the end of the cruise phase, the spacecraft will have a highly eccentric orbit with a close perihelion of 0.28 AU that is in resonance with Venus. This allows raising the inclination of the orbit using additional GAMS with Venus up to a maximum of  $28^\circ$  solar latitude during the nominal mission phase. Following the nominal mission, which will last for almost 4 years, the mission can be extended for 2.5 additional years, where the maximum solar latitude will increase to  $33^\circ$ . This orbit allows *Solar Orbiter* not only to observe the polar regions of the Sun (the spacecraft will spend up to 21 consecutive days at solar latitudes greater than  $30^\circ$ ), but also to observe the same region on the Sun for a longer time than possible from Earth (20 consecutive days for a viewing angle between  $\pm 75^\circ$ ).

### 1.4.3 Instruments

An important feature of the *Solar Orbiter* mission is the combination of remote-sensing and in-situ instruments. Measurements made by the in-situ instruments can easily be related to their source regions on the Sun. The remote-sensing instruments with high spatial resolution will all have the same field-of-view (FOV), allowing to connect measurements of the various atmospheric layers of the Sun. The instruments of *Solar Orbiter* can be grouped in three packages (for a detailed description of the individual instruments see Marsden & Müller 2011):

- Field Package: *Radio and Plasma Waves Instrument* (RPW) and *Magnetometer* (MAG).
- Particle Package: *Energetic Particle Detector* (EPD) and *Solar Wind Plasma Analyzer* (SWA).
- Solar remote sensing instrumentation: *Polarimetric and Helioseismic Imager* (PHI), *Extreme Ultraviolet Imager* (EUI), *Multi Element Telescope for Imaging and Spectroscopy* (METIS), *Solar Orbiter Heliospheric Imager* (SoloHI), *Spectral Imaging of the Coronal Environment* (SPICE) and *Spectrometer/Telescope for Imaging X-Rays* (STIX).

A key instrument is the *Polarimetric and Helioseismic Imager* (PHI). It will provide high-resolution maps of the continuum intensity, the Doppler velocity and the full magnetic field vector in the photosphere. Measuring the magnetic field in the photosphere is an important part of the mission since it is responsible for most of the magnetic activity in the upper layers of the atmosphere and the heliosphere. In addition, time-series of the Doppler velocity in the photosphere provided by PHI can be used for helioseismology and help probing the interior of the Sun.

### 1.4.4 The Polarimetric and Helioseismic Imager (PHI)

#### 1.4.4.1 Observables and Operation

PHI will perform a polarimetric and spectroscopic analysis of the Fe I 6173 Å line (the same as observed by the *Helioseismic and Magnetic Imager*, HMI, Schou et al. 2012). This line forms in the photosphere (formation height 100 – 400 km, Fleck et al. 2011) and is well suited for polarimetry (Norton et al. 2006). PHI will obtain images for four

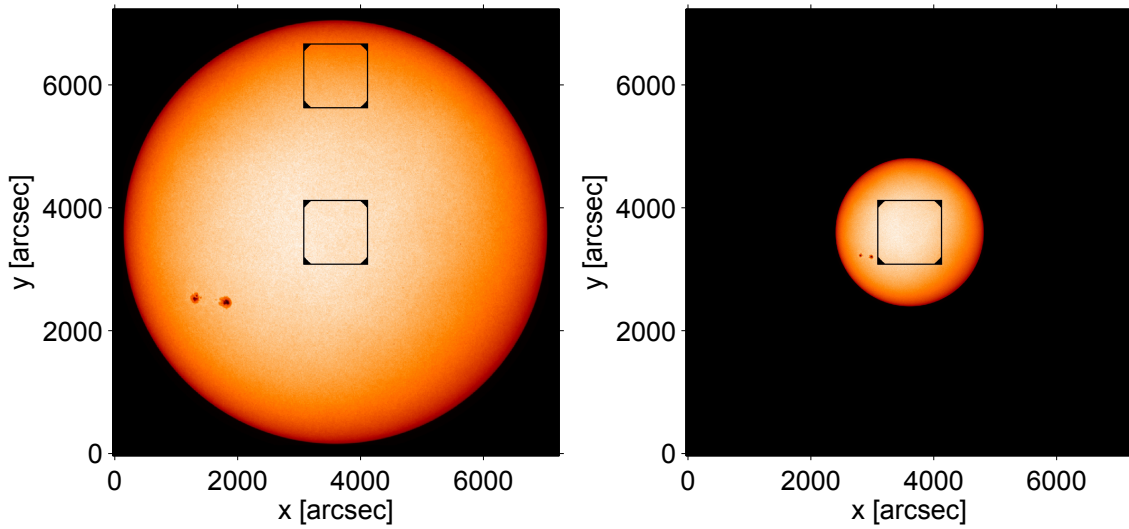


Figure 1.4: Field-of-view (FOV) of the two telescopes of PHI for a distance of 0.28 AU (*left*) and 0.8 AU (*right*) to the Sun. The *full images* correspond to the FOV of the FDT and the *small squares* to that of the HRT. PHI will not have a pointing mechanism of its own. Observations close to the limb with the HRT will be possible by changing the pointing of the spacecraft (*upper small square in the image on the left*). Courtesy of J. Hirzberger. Reprinted with permission.

polarization states at a few wavelength positions (current baseline: -400, -160, -80, 0, 80, and 160 mÅ from the line center) with a planned cadence of 60 s. These images will be used to perform an on-board inversion for the Doppler velocity and the full magnetic field vector. Due to telemetry constraints, only the inverted quantities plus the continuum intensity maps will be transferred to Earth.

The instrument consists of two telescopes that cannot be used simultaneously: A Full Disk Telescope (FDT) which can observe the full solar disk during all phases of the orbit, and a High Resolution Telescope (HRT) for observations of a smaller patch with high spatial resolution. Due to the changing distance to the Sun, the field-of-view and the spatial resolution of the two telescopes vary during the orbit (see Figure 1.4). The FDT will have a spatial resolution of about 9'' ( $\sim 1800$  km at disk center at perihelion) and the HRT will have a FOV of 16.8'' ( $\sim 200$  Mm at perihelion at disk center) and a spatial resolution of about 1'' ( $\sim 200$  km at disk center at perihelion).

There are two constraints for PHI, i.e. observing time and telemetry. Observing with PHI might interfere with observations by the in-situ instruments, and so, the current plan is to operate PHI and the other remote-sensing instruments only during three science windows per orbit. The science windows will have a length of ten days each and will be located at maximum northern and southern solar latitude and at perihelion. However, ten days of consecutive observations are too short for most of the science objectives involving helioseismology. It is currently under discussion to have an extended observing run dedicated to helioseismology. Further constraints on the observing time arise from the orbit. Direct observations of the poles, for example, are only possible during the excursions of the spacecraft to high solar latitudes. *Solar Orbiter* will be above 30° solar latitude for only up to 21 consecutive days. Long-term observations also require the spacecraft to

provide accurate thermal and pointing stability.

The second main challenge for observing with PHI is related to telemetry. The data transmission rate of *Solar Orbiter* will be extremely low. Currently, only 51 Gbits per science orbit are allocated to PHI. This corresponds to  $\sim 3.3$  kbps on average, much lower than what is needed for transmitting all raw data to Earth. Hence, observing with PHI requires an elaborated observing strategy. Especially helioseismology, with its need of a long time-series with high cadence, is affected by a low data rate. In order to reduce telemetry, the instrument will perform an on-board inversion for the line-of-sight velocity and the full magnetic field vector. In most cases, even the size of the inverted data will still exceed the available telemetry and additional lossy compression will have to be applied. It will only be possible to process the data when no observations are being performed, since the same computational resources will be used for data acquisition and processing. PHI will have 4 Tbits of flash memory. This will allow storing the data observed during the science windows and process it during the rest of the orbit. It can be transmitted whenever telemetry is available. Raw data will only be transmitted for calibration purposes.

In case of helioseismology, the LOS velocity can also be determined by using a simple algorithm, as it was successfully done by the *Michelson Doppler Imager* (MDI, Scherrer et al. 1995).

#### 1.4.4.2 Instrument Design

PHI is designed as a tunable narrow-band filtergraph, consisting of two telescopes, the FDT and the HRT. Both telescopes feed one filtergraph and one detector; therefore, they cannot be used simultaneously. The FDT is a refractive telescope with an aperture diameter of 17 mm, the HRT is designed as an off-axis Ritchey-Chrétien telescope with a decentered pupil of 140 mm diameter. In order to suppress pointing errors of the spacecraft (jitter, expected to be about  $0.5''$ ), the HRT includes an internal image stabilization system (ISS). A designated camera within the HRT measures the jitter of the spacecraft by cross-correlating images. The measured shift is then corrected by using a steerable mirror inside the HRT (down to  $\sim 0.1''$ ). The FDT is not equipped with an ISS since the expected spacecraft jitter is much lower than the pixel size of the FDT. See Figure 1.5 for a schematic drawing of PHI. The instrument is also described in Solanki et al. (2015).

*Solar Orbiter* will have a close orbit around the Sun. In order to protect the PHI instrument from the intense solar radiation, both telescopes will be equipped with special Heat Rejecting Entrance Windows (HREW) which are incorporated into the heat-shield assembly of the spacecraft. The HREWs have a transmittance of more than 80% at the science wavelength and almost perfect blocking from 200 nm to the far infrared. This significantly reduces the heat load entering the instrument, but has only minor influence of the high photometric and polarimetric accuracy of PHI.

Each telescope includes a Polarization Modulation Package (PMP) for the polarimetric analysis, consisting of two nematic liquid crystal variable retarders (LCVRs), followed by a linear polarizer as analyzer. It uses the same modulation scheme as the Imaging Magnetograph eXperiment (IMaX, Martínez Pillet et al. 2011) on the Sunrise balloon-borne observatory (Solanki et al. 2010, Barthol et al. 2011).

PHI will scan the Fe I 6173 Å line using a lithium niobate (LiNbO<sub>3</sub>) solid state etalon with a tuning range of about 2 Å (the line has a width of 120 mÅ). The transmission curve



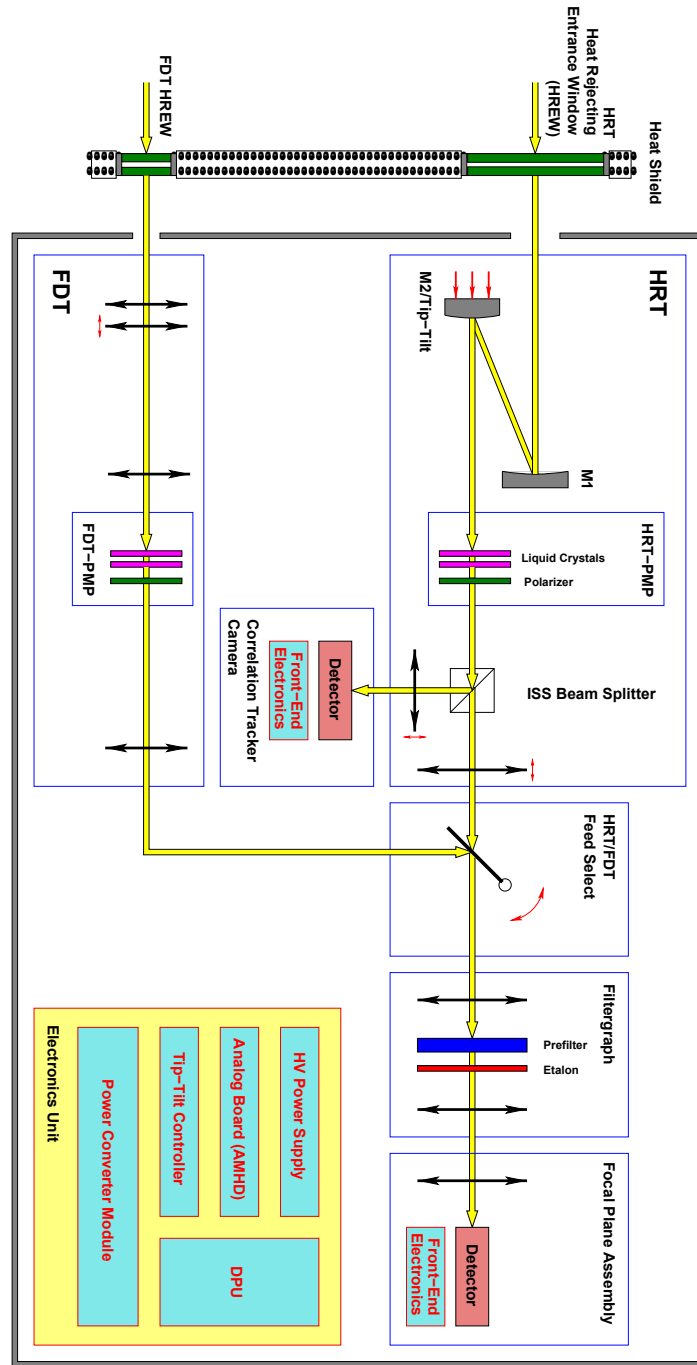


Figure 1.5: Functional block diagram of PHI. The instrument will consist of two telescopes, the Full Disk Telescope (FDT, *lower left part of the diagram*) and the High Resolution Telescope (HRT, *upper left part of the diagram*). Both telescopes have individual Polarization Modulation Packages (PMPs), but share the same filtergraph and one detector. More details are described in the text. Courtesy of J. Hirzberger. Reprinted with permission.

of the etalon exhibits peaks evenly distributed in wavelength. The correct peak is selected by two prefilters used in combination with the etalon. The wide-band prefilter has a Full Width at Half Maximum (FWHM) of  $100 \text{ \AA}$  and the narrow-band prefilter has a FWHM

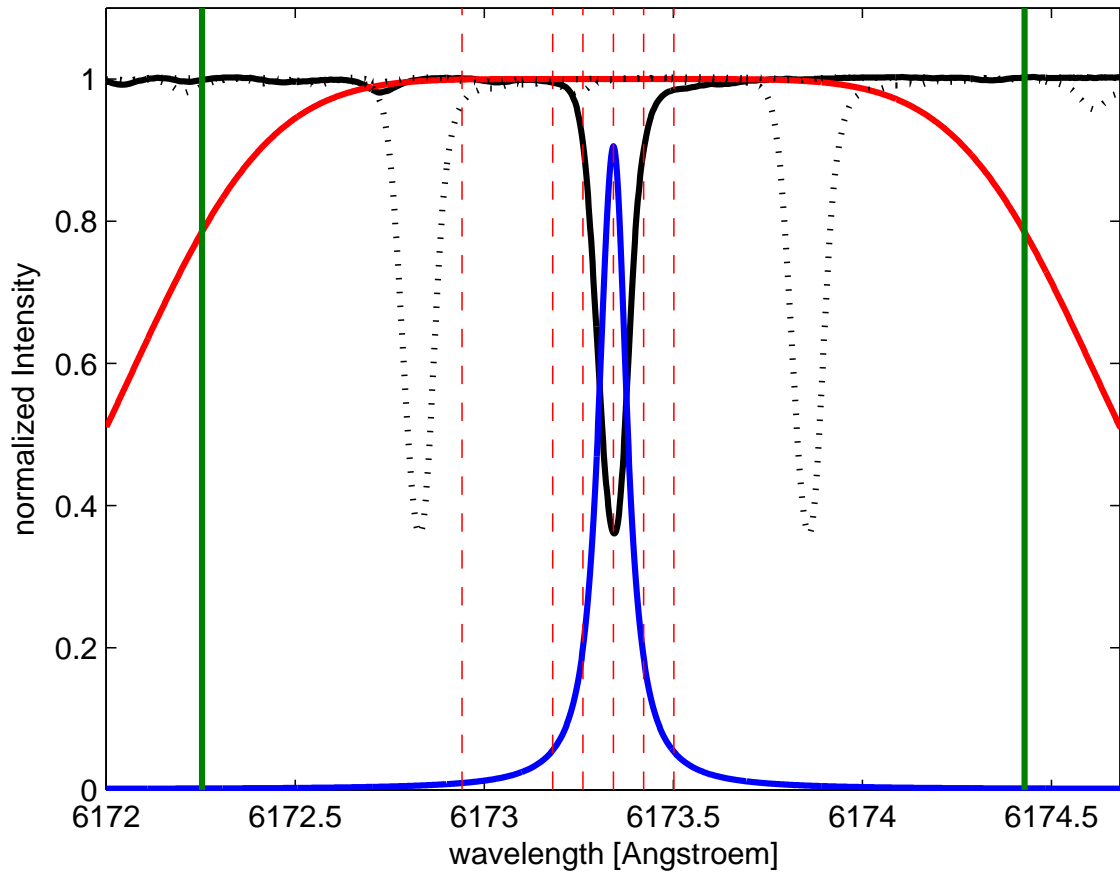


Figure 1.6: Schematic representation of the filter curves of the filtergraph used by PHI. The instrument will scan the Fe I 6173 Å line (*solid black line*) using three filters, the wide-band prefilter (not shown here), the order-selecting narrow-band prefilter (*red curve*) and a freely tunable etalon (*example profile shown by the blue curve*). The current baseline is to have the etalon scan the line at six wavelength positions (*vertical dotted lines*). The spacecraft will have a large orbital motion relative to the Sun (up to 25 km/s), causing the line to be Doppler-shifted significantly (up to  $\pm 0.5$  Å, *denoted by the dotted black lines*). This can be accounted for by the tuning range of the etalon (*given by the solid green lines*).

and 2.7 Å respectively (see Figure 1.6). The narrow-band prefilter is located in front of the etalon and selects its working order. The filter also transmits green light which is used for ground-based calibrations. This is removed by the wide-band prefilter, located behind the etalon. The tuning range of the etalon and the width of the narrow-band prefilter are sufficient to account for Dopplershifts of the line caused by the orbit (up to  $\pm 0.5$  Å).

PHI will use a  $2048 \times 2048$  pixel Active Pixel Sensor (APS) for imaging. It will obtain several images for every wavelength position and polarization state in order to increase the signal-to-noise ratio.

### 1.4.5 Science Goals Involving Helioseismology

*Solar Orbiter* offers new opportunities for helioseismology, but is also limited by constraints which are not relevant for current instruments. Instruments like HMI or GONG++ (Hill et al. 2003) observe the Sun from a fixed vantage point and have obtained almost continuous time-series for years. *Solar Orbiter*, on the other hand, will obtain only relatively short time-series. Hence, *Solar Orbiter* will not be able to perform more precise measurements than current instruments. The advantage of *Solar Orbiter* lies in its orbit. The spacecraft will observe the solar poles from a much lower viewing angle than possible now, allowing better measurements of large-scale and convective flows at high latitudes. In addition, the deep interior of the Sun could potentially be explored by using stereoscopic helioseismology. The science objectives of *Solar Orbiter* involving helioseismology are also discussed in Marsch et al. (2000), Gizon et al. (2001), Woch & Gizon (2007), Marsden & Müller (2011), and Löptien et al. (2014b).

#### 1.4.5.1 Near-Surface Flows at High Latitudes

**Large-Scale Flows** Meridional circulation and differential rotation are important features for understanding solar dynamics and the solar dynamo (Jiang et al. 2009, Dikpati & Gilman 2012), one of the primary science objectives of *Solar Orbiter*. While differential rotation is well measured at low solar latitudes, even in the deep interior (see e.g., review by Howe 2009), it is not well understood at solar latitudes above  $\sim 60^\circ$ . Global helioseismology, which is normally used for measuring rotation, is not sensitive near the poles because the global modes have only low amplitudes at high latitudes. The rotation profile near the poles is totally unknown; it might even exhibit a polar vortex (Gilman 1979), as already seen on planets (e.g., Fletcher et al. 2008). Solar rotation also varies with time, both within the solar cycle (torsional oscillations, Howard & Labonte 1980, Schou 1999, Howe et al. 2013) and from cycle to cycle (Hill et al. 2009). Repeated measurements of rotation at high latitudes with *Solar Orbiter* will especially help to characterize the high-latitude branch of the torsional oscillations.

Similarly, the meridional flow is not known at high latitudes. Observations at high latitudes will help determining, where and how the flow dives back into the Sun. This is a challenging objective, even at low latitudes, the meridional circulation has an amplitude of only 10 – 15 m/s (e.g., Zhao et al. 2013). It is expected to be even weaker at higher latitudes.

**Supergranulation** Convection in the Sun is still only poorly understood. In particular, the nature of supergranulation is still an issue of debate. Since many properties of supergranules vary in dependence on latitude, observations at high latitudes might be helpful for a better understanding. It is argued that supergranules prefer to align in the North-South direction at high solar latitudes (Nagashima et al. 2011). In addition, the influence of the Coriolis force on supergranulation (Gizon et al. 2010, Langfellner et al. 2014, 2015) and the wavelike properties of supergranulation reported in Gizon et al. (2003) and Schou (2003) clearly depend on latitude.

Supergranulation causes a strong signal in helioseismology and can easily be detected with the observing time and telemetry that are currently allocated to PHI.

### 1.4.5.2 Stereoscopic Helioseismology

**Deep Solar Interior** *Solar Orbiter* will be the first mission to test the concept of stereoscopic helioseismology. Based on this idea, the Sun is observed from two different vantage points using two different instruments, e.g., *Solar Orbiter* and HMI or GONG. In this way, a larger fraction of the surface of the Sun is covered, allowing to perform local helioseismology with waves penetrating the deep interior of the Sun.

Deep flows in the Sun are still poorly understood. While global helioseismology has successfully mapped differential rotation even in the deep interior, the shape of the deep meridional flow is still unclear. Recently, for example, Zhao et al. (2013) claimed that the meridional flow exhibits several cells with depth. Another important target for stereoscopic helioseismology is the tachocline at the border between the convection zone and the radiative interior. In many dynamo models, the strong shear in this layer generates the solar magnetic field. The dynamics of this layer, like the 1.3 year oscillations in the tachocline reported by Howe et al. (2000), is poorly understood. The nature of the deep convection in the Sun is also unclear. This refers to both the velocity amplitudes and the relevant spatial scales. The estimates of the velocity spectrum of deep convection by Hanasoge et al. (2012) are significantly lower than the prediction by current models. On the other hand, recent results by Greer et al. (2015) are in agreement with the simulations. New discoveries, such as the giant cells reported by Hathaway et al. (2013), could also be studied by PHI. In addition, stereoscopic helioseismology might also lead to the discovery of new phenomena that are not accessible with current methods.

Unfortunately, probing the deep interior of the Sun requires long observing times. Braun & Birch (2008) estimated that up to ten years would be required to measure a few m/s return meridional flow at the base of the convection zone when observing from a single vantage point. It is not known what noise level to expect when performing stereoscopic helioseismology. Since this method has not been used yet, only very preliminary estimates exist (Ruzmaikin & Lindsey 2003).

**Multi-Angle Observations** Combined observations with *Solar Orbiter* and another instrument can also be used to observe the same region on the Sun from two different vantage points. This would allow obtaining vector velocities and to study not only the relation between radial and horizontal flows in supergranules, but also the behavior of different components of the velocity vector of solar oscillations close to the surface. A similar study could be used to investigate the physics of MHD waves in active regions (e.g., Norton & Ulrich 2000).

Another application would be resolving systematic effects which depend on the distance to disk center, such as the shrinking-Sun effect in LCT or the phase shift of the solar oscillations in local helioseismology.

### 1.4.5.3 Other Science Objectives

**Active Regions and Sunspots** At perihelion, *Solar Orbiter* will have a low velocity relative to the Sun, allowing observations of the same region on the Sun for up to 20 days (viewing angle between  $\pm 75^\circ$ ). This is longer than what is currently possible and offers new opportunities for studying the evolution of active regions. However, this would be

possible with only a small number of active regions. Helioseismology of active regions typically uses an ensemble average of many active regions (e.g., Leka et al. 2013, Birch et al. 2013, Barnes et al. 2014).

When *Solar Orbiter* is on the far-side of the Sun, it could be used to calibrate far-side helioseismology (Lindsey & Braun 2000). Active regions present in *Solar Orbiter* data could be compared with the predictions made by another instrument using helioseismic holography.

**Low Resolution Observations** Observations of the Sun with low spatial resolution, e.g., from SOHO/VIRGO (Fröhlich et al. 1995) are regularly used for testing concepts used in asteroseismology (e.g., García 2009, Sato et al. 2010, Gizon et al. 2013). *Solar Orbiter* could contribute to this field by providing data from outside of the ecliptic. The relative amplitudes of stellar oscillation modes with same degree  $\ell$ , but different order  $m$  depends on the orientation of the rotation axis of the star to the observer. This effect is used for determining the inclination of stars (Gizon & Solanki 2003) and could be calibrated using *Solar Orbiter*. Observations outside the ecliptic could also be used for determining the shape of the Sun (e.g., Kuhn et al. 2012).

### 1.4.6 Observing Strategies for Helioseismology

Many of the science objectives presented in the previous section demand long observing times. The main sources of noise in local helioseismology are the realization noise caused by the stochastic excitation of the solar oscillations, convection, and, in case of near-surface flows, supergranulation. All these sources of noise are inherent to the Sun. The noise can only be reduced by long observing times or averaging over a large area. Supergranulation is the dominant source of noise for large-scale flows near the surface. Different supergranules ( $\sim 200$  m/s rms) are uncorrelated, and so, the noise scales with  $\sqrt{1/N}$ , where  $N$  is the number of supergranules. The number of supergranules that can be averaged over at a given latitude decreases with latitude, and so, the noise level increases with latitude (see Figure 1.7). For most science objectives, the length of the science windows (ten days) is not sufficient to achieve a satisfying signal-to-noise ratio. When observing the solar poles, the observing time is also limited by the orbit. The spacecraft will not be at solar latitudes greater than  $30^\circ$  for more than 21 consecutive days (for a launch in October 2018). Further challenges arise from the low telemetry rate of *Solar Orbiter*. Long observing runs are only possible if the data can be compressed very efficiently.

These challenges demand a very sophisticated observing strategy. For each science question, requirements for the noise level have to be made, determining the required observing time. Afterward, the influence of data compression on these measurements has to be evaluated. Compression adds not only noise to the data, but also a bias. Section 1.4.7 gives an overview of the different methods that could be used for compressing the data.

The observing strategy depends on the method used. Local helioseismology needs data with only modest or even low spatial resolution, depending on the oscillation modes used for the analysis, but it requires observing with a high cadence (at least 60 s). Flows near the surface could also be studied by using a feature tracking method like local correlation tracking of granulation (LCT, November & Simon 1988) or coherent structure tracking (CST, Rieutord et al. 2007, Tkaczuk et al. 2007, Roudier et al. 2013). These

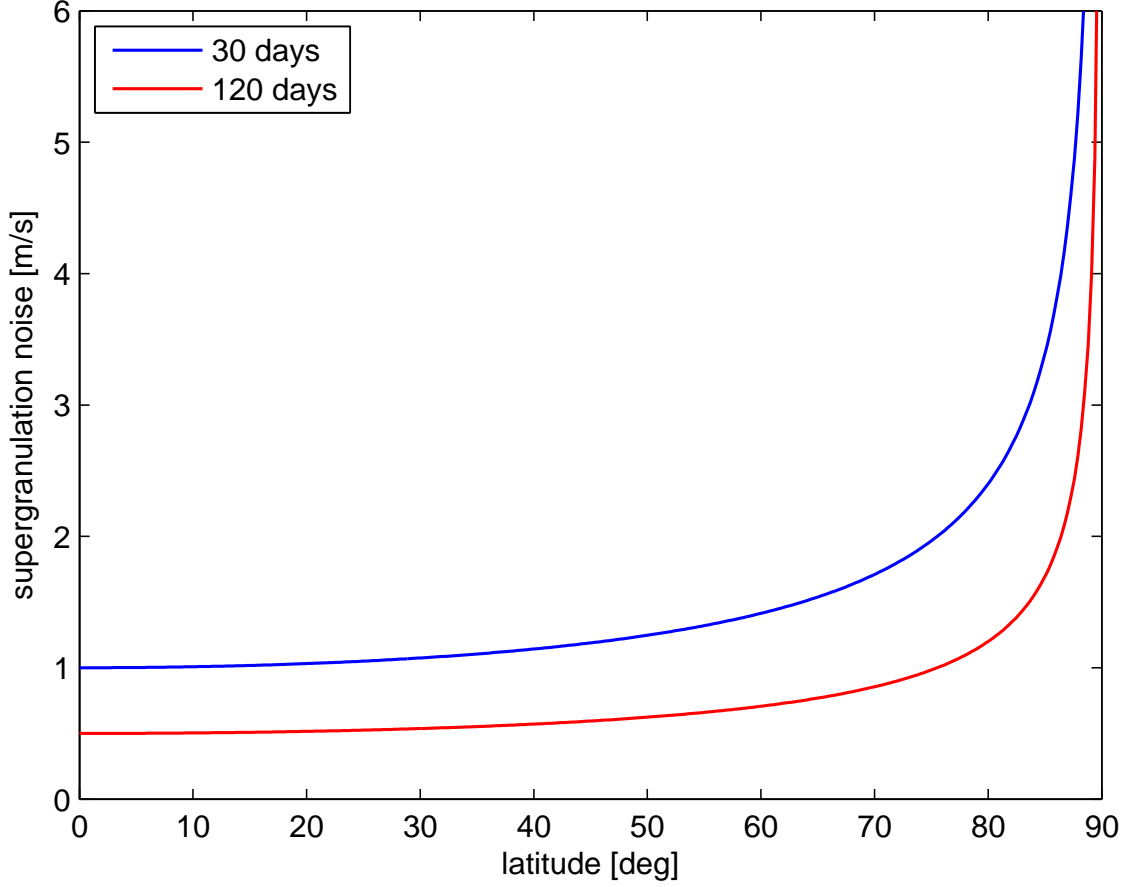


Figure 1.7: Estimate of the noise level originating from supergranulation (assuming that all longitudes are visible) for 30 days (*blue curve*) and 120 days (*red curve*) of continuous observations as a function of heliographic latitude,  $\lambda$ . Different supergranules are uncorrelated, and so, the noise scales with  $\sqrt{1/N}$ , where  $N$  is the number of supergranules. The number of supergranules that can be averaged over at a given latitude, decreases with latitude. So, the noise level is proportional to  $(\cos \lambda)^{-1/2}$ , where  $\cos \lambda$  is the ratio of the number of supergranules between the equator and latitude  $\lambda$ . For this plot, a noise level of 1 m/s at the equator for 30 days of data (Figure 6 of Gizon et al. 2001) and a supergranulation granulation lifetime of 24 h were assumed. Supergranulation noise affects all methods that measure flows at or near the surface of the Sun.

need a high spatial resolution, but it might be possible to reduce the cadence significantly by obtaining pairs of images, where the pairs are separated by minutes or even hours. Another option for probing large-scale flows would be to track supergranules as done by Hathaway et al. (2013) or to track magnetic features. The required spatial resolution defines which telescope has to be used: low resolution observations could be made with the FDT, high resolution data requires the HRT. Table 1.1 gives an overview on the parameters that need to be decided on for an observing strategy.

Table 1.2 summarizes first estimates for observing strategies for the individual science goals discussed in the previous section. I assumed a file size of 5 bits per pixel for computing the telemetry estimates. See Löptien et al. (2014b) for more information

Table 1.1: Parameters to be selected when planning an observing strategy.

Parameter	Options
Method	e.g., time-distance helioseismology or LCT
Observables	$v_{\text{los}}$ , $I_c$ , $B_{\text{los}}$ , depending on selected method
Cadence	set by method and telemetry
Duration	determines noise level, set by telemetry and orbit
Telescope	FDT, HRT, depending on required spatial resolution
Spatial sampling	depending on required spatial resolution
Compression	as little as possible, depending on telemetry
Vantage point	High inclination to study polar regions; Full range of Earth-Sun-spacecraft angles for stereoscopy

on these estimates. Many science objectives are not within the baseline allocations of telemetry and observing time. The observing time is fixed by the required noise level, but it might be possible to reduce the required telemetry by using different methods of data compression (see Section 1.4.7).

All methods for measuring flows suffer from systematic errors that depend on the distance to disk center. Current attempts of measuring the meridional flow by time-distance helioseismology, for example, are affected by a systematic error causing a phase shift of the travel times with latitude (Zhao et al. 2012). This is probably caused by a height-dependent asymmetry of granular flows (Baldner & Schou 2012). In local correlation tracking, an artifact looking like a converging flow towards disk center, also called shrinking-Sun effect (Lisle & Toomre 2004), interferes with measurements far from disk center. One of the advantages of *Solar Orbiter* is that due to inclined orbit the influence of systematic errors at high latitudes will be smaller than for current instruments. However, even for an inclination of  $30^\circ$  solar latitude (as will be reached by the spacecraft), the solar pole is  $60^\circ$  away from disk center and systematic errors are still not negligible. The viewing geometry also changes with time, due to the orbit. This leads to time-dependent systematic errors in the data.

Further systematic errors might arise from the instrument. Due to the strongly varying distance to the Sun, the thermal environment of PHI changes. This might affect e.g., the point spread function. In addition, the pointing of PHI must be well known during all phases of the orbit. When observing with the HRT, the limb is not necessarily within the image and cannot be used for determining the position on the Sun. Similarly, the roll angle of the spacecraft must be known. This requires an extremely good calibration of the instrument.

### 1.4.7 Data Compression

Testing compression is of great importance for *Solar Orbiter*. The vast amount of data obtained for helioseismology have to be transferred to Earth with minimal influence on the scientific output. This section gives a brief overview on the different methods that could be used for compressing the data. The method of choice depends upon the required amount

Table 1.2: Table of helioseismology science objectives. Some science goals exceed the currently allocated observing time or telemetry (highlighted in red, green colors indicate that the parameter is within the allocation). The table is continued on the next page. The table is reprinted from Löptien et al. (2014b). I contributed to developing the observing strategies and generated the table.

Science target	# Spatial points	Observing time	Observables	Approx. telemetry (5 bits/observable)
<b>Near-surface rotation, meridional circulation, and solar-cycle variations at high latitudes</b>				
- Helioseismology	512 × 512	30 days	$V_{\text{los}}$ every 60 s	60 Gbit
- Solar-cycle variations from helioseismology	512 × 512	4 × 30 days, 2 years apart	$V_{\text{los}}$ every 60 s	4 × 60 Gbit
- Meridional circulation to 3 m/s at 75° (see Figure 1.7)	512 × 512	100+ days	$V_{\text{los}}, I_c \& B_{\text{los}}$	method dependent
- Granulation and magnetic-feature tracking	2048 × 2048	30 days	$I_c \& B_{\text{los}}$ , two consecutive images every 8 h	8 Gbit
- Supergranulation tracking	512 × 512	30 days	$V_{\text{los}}$ every 60 min	1 Gbit
<b>Deep and large-scale solar dynamics</b>				
- MDI-like medium- $l$ program	128 × 128	continuous	$V_{\text{los}}$ every 60 s	40 Gbit/year
- Stereoscopic helioseismology (PHI + other instrument)	128 × 128	continuous	$V_{\text{los}}$ every 60 s	40 Gbit/year
<b>Convection at high latitudes</b>				
- Helioseismology	1024 × 1024	7 days	$V_{\text{los}}$ every 60 s	50 Gbit
- Feature tracking	2048 × 2048	7 days	$I_c \& B_{\text{los}}$ , two consecutive images every 8 h	2 Gbit



<b>Deep convection and giant cells</b>				
- Helioseismology	128 × 128	continuous	$V_{\text{los}}$ every 60 s	40 Gbit/year
- Feature tracking	512 × 512	4 × 60 days	$V_{\text{los}}$ every 60 min	4 × 2 Gbit
<b>Active regions and sunspots</b>				
- Active region flows & structure	512 × 512	20 days	$V_{\text{los}}$ every 60 s	40 Gbit
- Sunspot oscillations	1024 × 1024	2 days	$V_{\text{los}}, I_c$ & $\vec{B}$ every 60 s	80 Gbit
- Calibration far-side helioseismology	128 × 128	5 × 2 days	$B_{\text{los}}, I_c$ & $V_{\text{los}}$ every 60 s	5 × 0.3 Gbit
<b>Physics of oscillations (stereoscopic obs.)</b>				
- Effect of granulation on oscillations	2048 × 256	1 day	6 filtergrams every 60 s	20 Gbit
- Two components of velocity	512 × 512	10 days	$V_{\text{los}}$ every 60 s	20 Gbit
- Magnetic oscillations	2048 × 2048	1 day	$V_{\text{los}}, I_c, B_{\text{los}}$ every 60 s & $\vec{B}$ at max. cadence	100 Gbit
<b>Low resolution observations</b>				
- LOI-like observations (solar-cycle variations, active longitudes)	4 × 4 for $V_{\text{los}}$ & $I_c$ , 32 × 32 for $B_{\text{los}}$	continuous	$V_{\text{los}}$ & $I_c$ every 60 s & $B_{\text{los}}$ once per day	0.1 Gbit/year
- Shape of the Sun	10 × 6000	Every few months	$I_c$ at 12 angles during rolls	4 Mbit for one roll

of compression. If only a modest compression rate is necessary, a lossless compression algorithm might already be sufficient. Lossy compression involves a trade-off between the compression efficiency and the influence on the science derived from the data. Lossy compression adds noise and causes a bias, which depend on the compression method, the data analysis method, and the science goal. This needs to be tested carefully before implementing an algorithm. Besides, it has to be taken into account that the complexity of a compression algorithm is limited by the computational resources available on-board.

### 1.4.7.1 Lossless Compression

When uncompressed data are stored, e.g., a gray-scale image, the value of each pixel is stored using a fixed precision. Hence, the same amount of data is required for each pixel, for example 32 bits per pixel for single precision. Lossless compression methods reduce the file size by using statistical properties of the data. Values with a high probability are encoded using fewer bits per pixel than less probable ones. The details are different for the various existing lossless compression algorithms.

An example for a lossless compression algorithm is Huffman compression (Huffman 1952) which is visualized in Figure 1.8. The two symbols with the lowest probabilities are combined and their probabilities summed up. One bit is used for distinguishing between these symbols. In the next step, the combination of these symbols is combined with the symbol with the next higher probability. This results in a so-called Huffman tree.

### 1.4.7.2 Quantization

Quantization means dividing the data and rounding to the nearest integer. This reduces the number  $n$  of the possible values of the data and also the number of bits per pixel needed for storing the data. The lower  $n$ , the stronger the influence of the compression. For large  $n$ , the error introduced by the compression is almost uncorrelated between adjacent pixels and corresponds to adding white noise to the data. For small  $n$ , the compression causes sharp edges in the data. For  $n = 2$ , only positive and negative variations relative to the mean can be distinguished.

Quantization can be combined with a lossless compression algorithm.

### 1.4.7.3 Spatial Transformations

Solar oscillations act on distinct temporal and spatial scales. The compression can make use of this property by applying some spatial or temporal transformation to the data and transmitting only specific coefficients. An example for such a method is JPEG compression (Wallace 1992).

When using these methods, the spherical geometry of the Sun has to be taken into account. Due to foreshortening, the spatial resolution (in km/pixel) decreases towards the limb in center-to-limb direction, but it stays constant parallel to the limb. The ideal spatial transformation would be a decomposition into individual spherical harmonics. In combination with a temporal Fourier transform, this would allow transmitting the relevant data for the solar oscillations without any loss of important information.

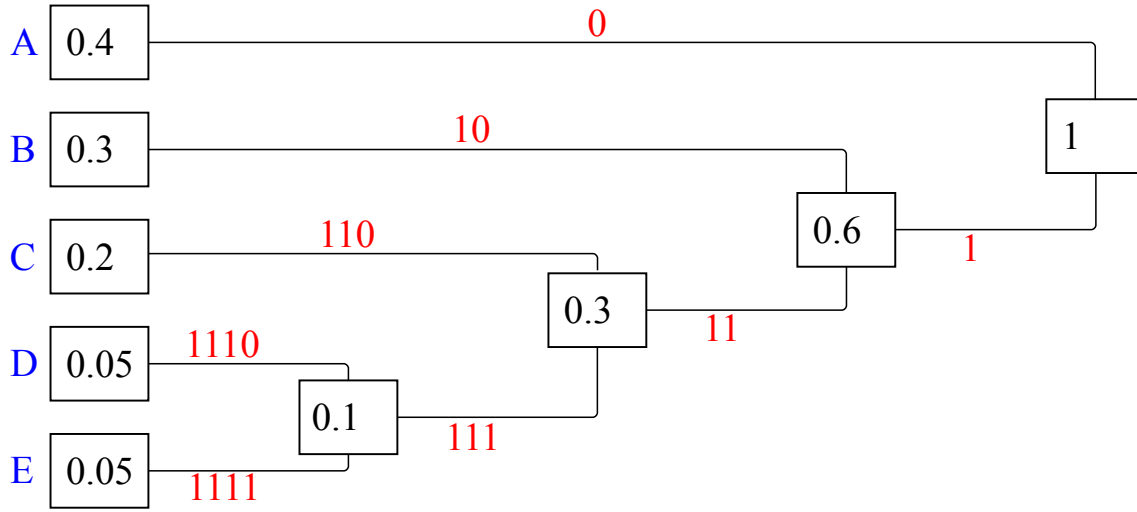


Figure 1.8: Example of Huffman compression. The data to be compressed are represented by five symbols (A,B,C,D,E, shown in blue) appearing with certain probabilities ( $p(A) = 0.4$ ,  $p(B) = 0.3$ ,  $p(C) = 0.2$ ,  $p(D) = 0.05$ ,  $p(E) = 0.05$ ). Huffman coding compresses these data by changing the number of bits for coding a symbol depending on its probability: The lower the probability, the more bits are used. Starting from the lowest probability, the symbols are combined into pairs and one bit is used for distinguishing between them. In this example, D and E have the lowest probabilities. In the next step, D and E are treated as one symbol with a joint probability of 0.1. One bit is used to separate D/E and the symbol with the next higher probability (symbol C). This procedure is continued until all symbols are incorporated and the total probability has summed up to one. Depending on the probability of a distinct symbol, between one and four bits are necessary for encoding it in this example. On average, two bits per symbol are needed here, whereas three bits are required in case of applying no compression.

#### 1.4.7.4 Binning, Subsampling, and Cropping

These methods are the simplest ones for compressing data. Cropping the image means transmitting only the parts of the CCD images that are needed for the scientific analysis. If a high spatial resolution is not required, binning or subsampling the images is also an option. For helioseismology, however, the data should be smoothed with a low-pass filter, e.g., by convolving them with a Gaussian, before subsampling the data. Otherwise, the reduction of the resolution can lead to aliasing in the power spectrum. Subsampling the images can reduce the file size significantly and this method can easily be combined with one of the other methods presented in this section.

These methods were used by MDI. In the medium- $\ell$  program, the images were cropped at  $0.9R_{\odot}$ , smoothed with a 2D Gaussian, and then subsampled (Kosovichev et al. 1996, 1997).

#### 1.4.7.5 On-board Analysis

Instead of transmitting compressed images and doing the scientific analysis on ground, some simple analysis could in principle also be done on-board. In that case, only the results of the analysis would have to be transferred to Earth. A possible example for this

would be local correlation tracking. Running an LCT code on-board is easy to implement and would reduce the required telemetry significantly. Instead of a long time-series of continuum intensity maps or magnetograms, only the derived flow map would have to be transferred. The disadvantage of this method is that the raw data used for the analysis would have to be calibrated extremely well. The geometry of the data would have to be well known.

## 2 Simulating PHI Time-Series for Helioseismology

This chapter appeared as Sections 6 and 7 in the paper *Helioseismology with Solar Orbiter* (Space Science Reviews, 2014, in press) by B. Löptien, A. C. Birch, L. Gizon, J. Schou, et al. I designed, carried out, and wrote the work presented in this chapter, except Section 2.2.3, which was done in collaboration with J. Blanco Rodríguez (GACE/IPL, Universidad de Valencia, Spain).

### 2.1 Introduction to Chapter 2

We are now at the stage where a detailed strategy has to be developed to maximize the helioseismology output of the mission given the various limitations imposed by the mission (e.g., challenging orbit, and an expected telemetry allocation of 51 Gbit per orbit). In this section we present synthetic data with the same properties as expected from the High Resolution Telescope of the PHI instrument on *Solar Orbiter* and begin characterizing the properties of the data for helioseismic studies (e.g., the expected power spectra). Starting from realistic radiative MHD simulations computed with the STAGGER code (Stein & Nordlund 2000), computing line profiles with the SPINOR code (Frutiger et al. 2000), and simulating the PHI instrument using SOPHISM, we have generated a time-series of synthetic Dopplergrams. These Dopplergrams are models for what should be available onboard the *Solar Orbiter* satellite.

### 2.2 Steps in the Generation of Synthetic Data

#### 2.2.1 Simulations of Solar Convection

We start from simulations of solar surface convection for the quiet Sun computed with the STAGGER code. These simulations exhibit solar oscillations and have previously been used in helioseismic studies, mostly for analyzing the excitation mechanism of solar oscillations (e.g., Stein & Nordlund 2001, Nordlund & Stein 2001, Samadi et al. 2003, Stein et al. 2004) but also for testing methods used in helioseismology (e.g., Zhao et al. 2007, Georgobiani et al. 2007, Braun et al. 2007, Couvidat & Birch 2009) or for modeling helioseismic observations (e.g., Baldner & Schou 2012).

We use a simulation run with a size of  $96 \times 96 \times 20$  Mm, corresponding to  $2016 \times 2016 \times 500$  grid points. The horizontal resolution is constant (about 48 km) and the vertical resolution varies with height between 12 and 79 km. We analyze 359 minutes of the simulation, for which we have snapshots of the entire state of the system with a cadence of one minute corresponding to the planned cadence of PHI. We assume in this work that variations on shorter timescales have a negligible effect. We plan to test this assumption using a short run of the simulations with a high cadence.

In order to reduce computation time we analyze a  $48 \times 48$  Mm sub-domain from the simulations. This is a small patch of the field of view of the HRT of PHI (16.8', corresponding to  $\sim 200$  Mm at perihelion) but it is sufficient for studying the solar oscillations in the synthetic data.

### 2.2.2 Computation of Line Profiles

We synthesize line profiles for the Fe I 6173 Å line for every single pixel of the simulations using the SPINOR code with atomic parameters taken from the Kurucz atomic database ( $\log gf = -2.880$ , Fuhr et al. 1988) and the iron abundance ( $A_{\text{Fe}} = 7.43$ ) from Bellot Rubio & Borrero (2002). Note that SPINOR allows simulations of observations at heliocentric angles  $\rho > 0$  by a synthesis of spectra obtained from inclined ray paths.

### 2.2.3 Simulation Tool: SOPHISM

SOPHISM is a software simulator aimed at a full representation of the PHI instrument (both hardware and software) and is applicable to both telescopes of the instrument. Starting from 2D maps of line profiles for the Fe I 6173 Å line computed from MHD simulations, SOPHISM generates synthetic 2D maps of the observables that will be measured by PHI. This allows estimates of the performance of PHI, which helps optimize the instrument design and the onboard processing. The hardware part of the simulations takes into account all the elements affecting solar light from when it enters the telescope (including some perturbations due to the spacecraft) up to the registering of this light on the detector. The software part deals with the data pipeline and processing that takes place onboard.

The simulator is programmed in the Interactive Data Language<sup>1</sup> (IDL) in a modular structure, each module deals with one aspect of the instrument. The modules are mostly independent from each other and can individually be enabled or disabled. The simulation runs can be saved at each step and made available for subsequent calculations. Also, the code is very flexible, with various parameters that can be modified. Presently, the modules and effects covered are the following:

- **Input:** this module prepares the input data to be used in the simulation run. If needed, temporal interpolation is performed, as well as spatial operations such as replication of the FOV (for simulation data with periodic boundary conditions) or, considering the Sun-spacecraft distance, scaling from the original spatial resolution to that of the detector.
- **Jitter:** this module represents the vibrations induced by the spacecraft, including also the correction by the ISS. A random shift of the FOV is generated and then

---

<sup>1</sup>IDL is a product of EXELIS Visual Information Solutions, <http://www.exelisvis.com/>

filtered in frequency according to different possibilities, including a jitter model similar to the *Hinode* spacecraft (Katsukawa et al. 2010). Next, the resulting jitter is diminished by means of the ISS attenuation function.

- **Polarization:** the polarization modulation of the incoming light is parametrized in this module. This comprises the Mueller matrix of the system and the liquid crystal variable retarders (LCVRs) settings, such as orientation and retardances, to produce the desired modulation of the data.
- **Spectral Profile:** the spectral transmission profile of the instrument is calculated here at the user-defined wavelength positions. The resulting transmission curve, a combination of the prefilter and etalon transmissions, is convolved with the data. An example of a simulated spectral transmission profile can be seen in Figure 1.6.
- **Optical Aberrations:** the MTF of the system with its different aberrations (e.g., defocus, astigmatism, coma,...) is characterized and convolved with the data.
- **Pupil apodization:** since the etalon is placed in a telecentric mounting, close to focus, the light is converging over it and enters as a cone. This produces the so-called pupil apodization (Beckers 1998, von der Lühe & Kentischer 2000), which results in a radial gradient of intensity and phase over the pupil. This effect has consequences for the spatial resolution and spectral transmission which are taken into account in this module.
- **Focal Plane:** all the detector effects are included in this module (e.g., dark current, readout noise, flat-field, shutter, etc.).
- **Accumulation:** a given number of exposures are added directly on the instrument in order to achieve a better signal-to-noise ratio.
- **Demodulation:** the Stokes vector is recovered here using the demodulation matrices calculated at the polarization module.
- **Inversion:** the MILOS inversion code (Orozco Suárez & Del Toro Iniesta 2007), based on a Milne-Eddington approximation, is used to retrieve the full vector magnetic field and LOS velocities maps from the demodulated Stokes vector obtained by the instrument.

Figures 2.1 and 2.2 shows an example of a simulation run. Although in the following part of this paper we only use Stokes  $I$ , we present here results for the magnetic Stokes parameters as well. The left panels correspond to the input data (MURaM MHD simulations with an average magnetic field of 50 G, Vögler et al. 2005), the right panels to the output of SOPHISM. The images presented here are intensity at line center and in the continuum, and  $\sqrt{Q^2 + U^2}/I$  and  $V/I$  at  $-80 \text{ mÅ}$  in the line wing, corresponding to the degree of linear and circular polarization, respectively. The simulations represent an observation at disk center from perihelion (0.28 AU) and have a size of  $12 \times 12 \text{ Mm}$  (the original simulations have a size of  $6 \times 6 \text{ Mm}$ , but are replicated in the horizontal spatial dimensions in order to increase the FOV, taking advantage of its periodic boundary conditions). The simulations presented in this section have the following main characteristics: jitter

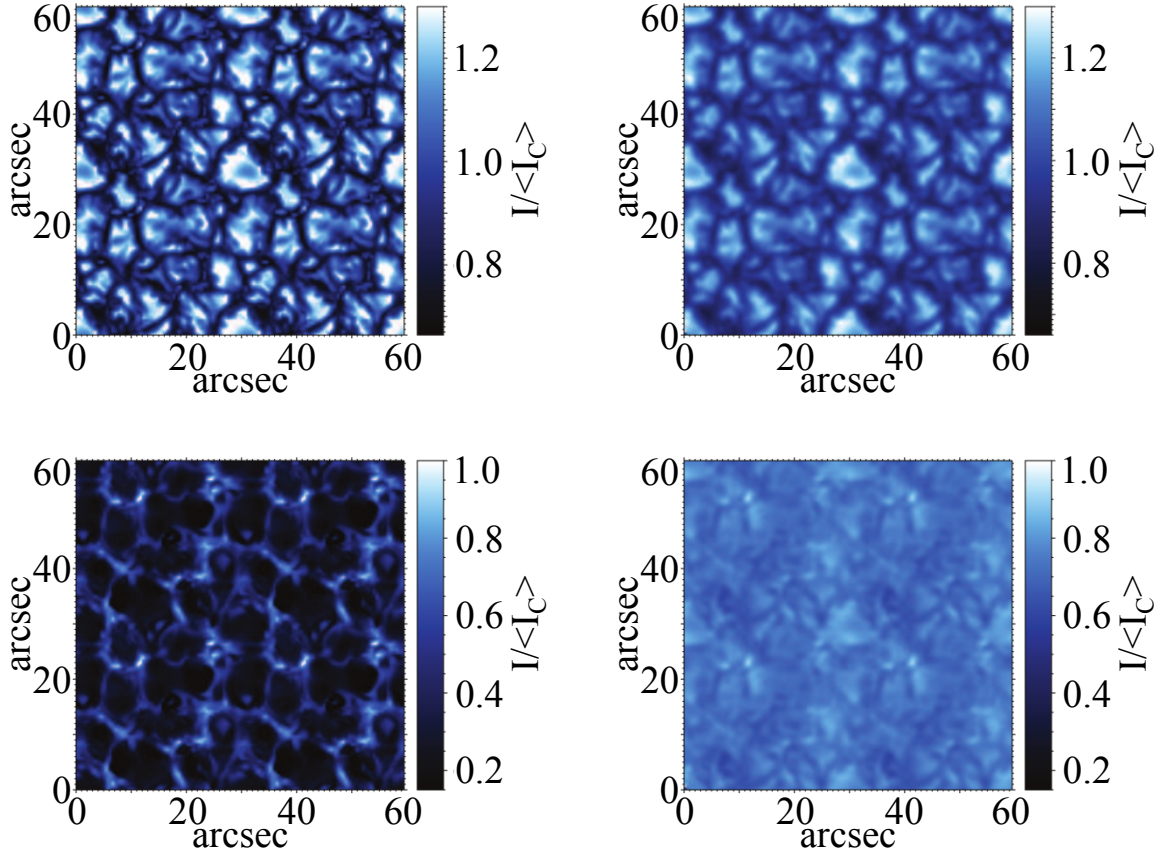


Figure 2.1: Example of the influence of SOPHISM on intensity. The *left column* of panels shows the original data from an MHD simulation, the *right column* of panels show the results after running SOPHISM. *Top*: Stokes  $I$  at  $-400 \text{ m}\text{\AA}$  from line center, *bottom*: Stokes  $I$  at line center. The configuration of SOPHISM that was used here and further details about the images are described in the text.

with an RMS normalized to  $0.5''$  and attenuated by the ISS; polarization modulation with identity Mueller matrix and ideal modulation; a prefilter with a FWHM of  $3 \text{ \AA}$  and the etalon at the six positions given in Section 1.4.4 with a resulting transmission FWHM of  $\sim 90 \text{ m}\text{\AA}$  along with pupil apodization considerations; an aberrated wavefront normalized to  $\lambda/10$ ; dark current and photon and readout noises; and 12 accumulations.

In the intensity images presented in Figures 2.1 and 2.2, the most evident effect of the simulated instrument is a loss of spatial and spectral resolution. Because of the lower signal of the magnetic Stokes parameters  $Q$ ,  $U$ , and  $V_{\text{los}}$ , other effects become apparent in the linear and circular polarization images, mainly crosstalk from Stokes  $I$  due to spacecraft jitter and noise components. The random parts of the simulation, like the jitter generation, noises, etc., are produced in every run. So, a different simulation run, even with the same settings, will yield different results. This is especially noticeable with the jitter crosstalk in the  $Q$ ,  $U$ ,  $V_{\text{los}}$  parameters, which may show different contributions in the same parameter at the same spectral position because of larger or smaller jitter shifts coinciding in the same data product.



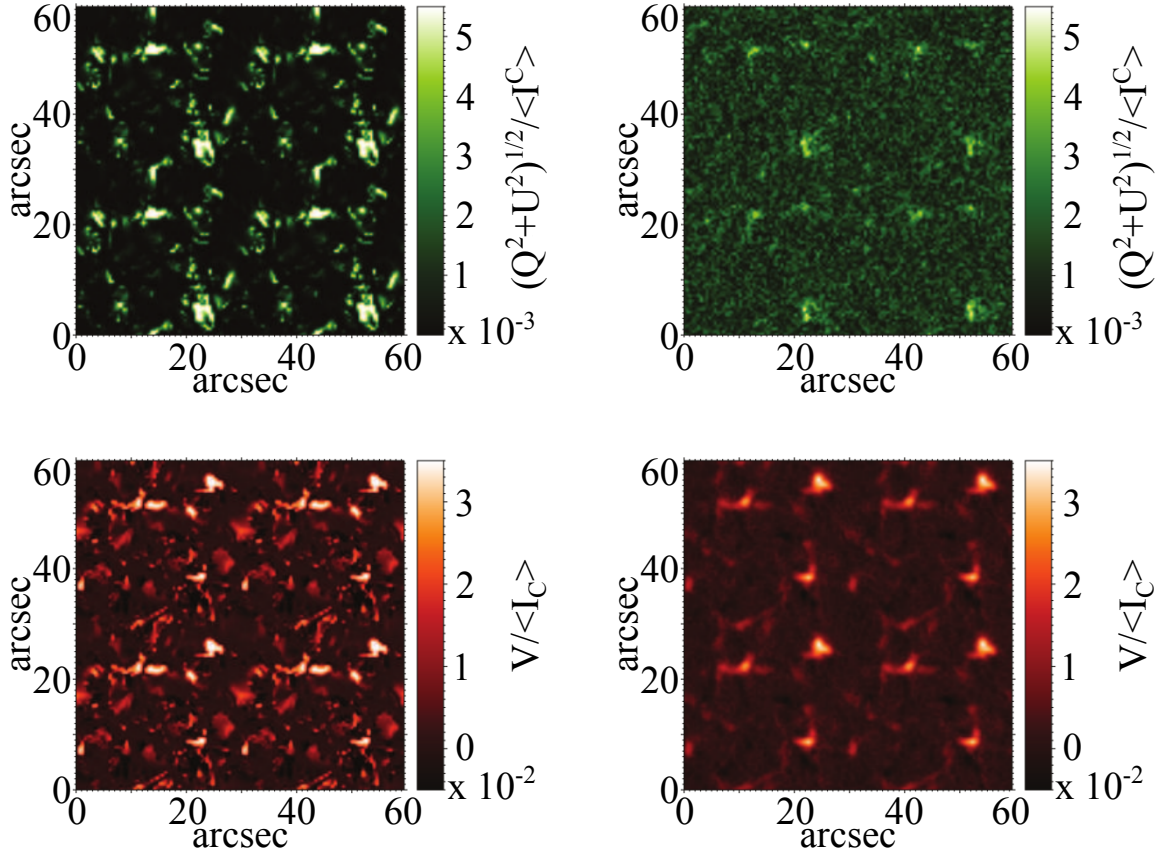


Figure 2.2: Example of the influence of SOPHISM on polarization. The *left column* of panels shows the original data from an MHD simulation, the *right column* of panels show the results after running SOPHISM. *Top*:  $\sqrt{Q^2 + U^2}/I$  at  $-80$  mÅ *bottom*:  $V/I$  at  $-80$  mÅ from line center, corresponding to the degree of linear and circular polarization. The configuration of SOPHISM that was used here and further details about the images are described in the text.

The output of SOPHISM depends significantly on the initial settings mentioned above. Although the configuration used for the results shown in Figures 2.1 and 2.2 correspond to the present assumptions for the behavior of PHI, these simulations will not be identical with the actual instrument. The default settings and characteristics of most elements are taken from other instruments or calculated from a theoretical point of view. For example, the frequency filter for the spacecraft jitter is taken from *Hinode* and the prefilter curve is computed theoretically. As the design of PHI proceeds, more precise settings can be used in the simulator, including parameters derived from measurements in the lab. The development of the code will continue during the following years to represent further realistic characteristics of the instrument, such as temperature dependencies, aging (radiation effects, inefficiencies,...), FOV dependent Mueller matrices, etc.

## 2.2.4 Using SOPHISM

Since we want to analyze a relatively long time-series, we configure SOPHISM in a way that minimizes the computation time while still covering all relevant effects for helioseis-

mology. One of the computationally most demanding parts of SOPHISM is the polarimetry. Polarimetry is not of great importance for helioseismology, so we do not include it in the following. We further reduce the computation time by decreasing the number of simulated exposures. When scanning the Fe I 6173 Å line, PHI will make several exposures for each of the wavelength positions over a total time of 60 s. We assume observations at six wavelength positions, located at  $-400$ ,  $-160$ ,  $-80$ ,  $0$ ,  $80$ , and  $160$  mÅ relative to the line center. During this time, the Sun evolves, causing small differences between the individual exposures. This means that the order in which the wavelengths are observed affects the resulting image. The STAGGER simulation data we use are sampled at a cadence of 60 s, so we can compute only one exposure. The number of exposures and the order in which they are taken are also important for the impact of spacecraft jitter and photon noise. So, in order to get realistic results when computing only one exposure, in this section we treat the jitter and photon noise in a slightly different way than the standard SOPHISM.

### 2.2.5 Spacecraft Jitter

The exact behavior of the spacecraft jitter of *Solar Orbiter* is not known yet. When modeling the jitter, we follow the spacecraft requirement that the RMS of the jitter with frequencies above 0.1 Hz must not be higher than 0.5'' (or 0.1'' when the ISS is turned on). We use a model for the jitter that depends on frequency with a power spectral density similar to that of the *Hinode* spacecraft. For frequencies higher than 0.1 Hz, we model the power spectral density with a power law fit of the *Hinode* jitter curve (see left part of Figure 2.3) normalized to an RMS of 0.5''. We are not aware of specifications for the jitter at lower frequencies. Here we assume a constant power spectral density of the jitter, corresponding to the power at  $\nu = 0.1$  Hz. This low frequency part increases the RMS from 0.5'' to 0.7''. We also model a jitter curve when the ISS is activated by dividing the power spectral density of the jitter by a modeled attenuation curve of the ISS (black curve in Figure 2.3). The attenuation is strongest at low frequencies and reduces the RMS from 0.7'' to 0.03'', which is much less than the pixel size (0.5''). We express the jitter in the Fourier domain by using a fixed amplitude for each frequency (derived from the power spectral density) and a random phase. We generate a time-series of the jitter by taking the inverse Fourier transformation.

The influence of the jitter depends on the order in which the filtergrams are measured. We assume the jitter to act on all wavelengths in the same manner by using the same time-series for all wavelengths. Since we compute only one exposure, we do not model the jitter by shifting individual images. Instead, we convolve the image with a distribution corresponding to the accumulated shifts introduced by the jitter within the PHI cadence. The right part of Figure 2.3 shows an example for this distribution. Since the jitter is derived from *Hinode*, it does not correspond to the jitter of the final *Solar Orbiter* spacecraft. However, as will be shown in Section 2.4, the influence of the jitter on the power spectrum is very small. Hence, we do not expect that our conclusions are affected by selecting a specific model for the jitter.

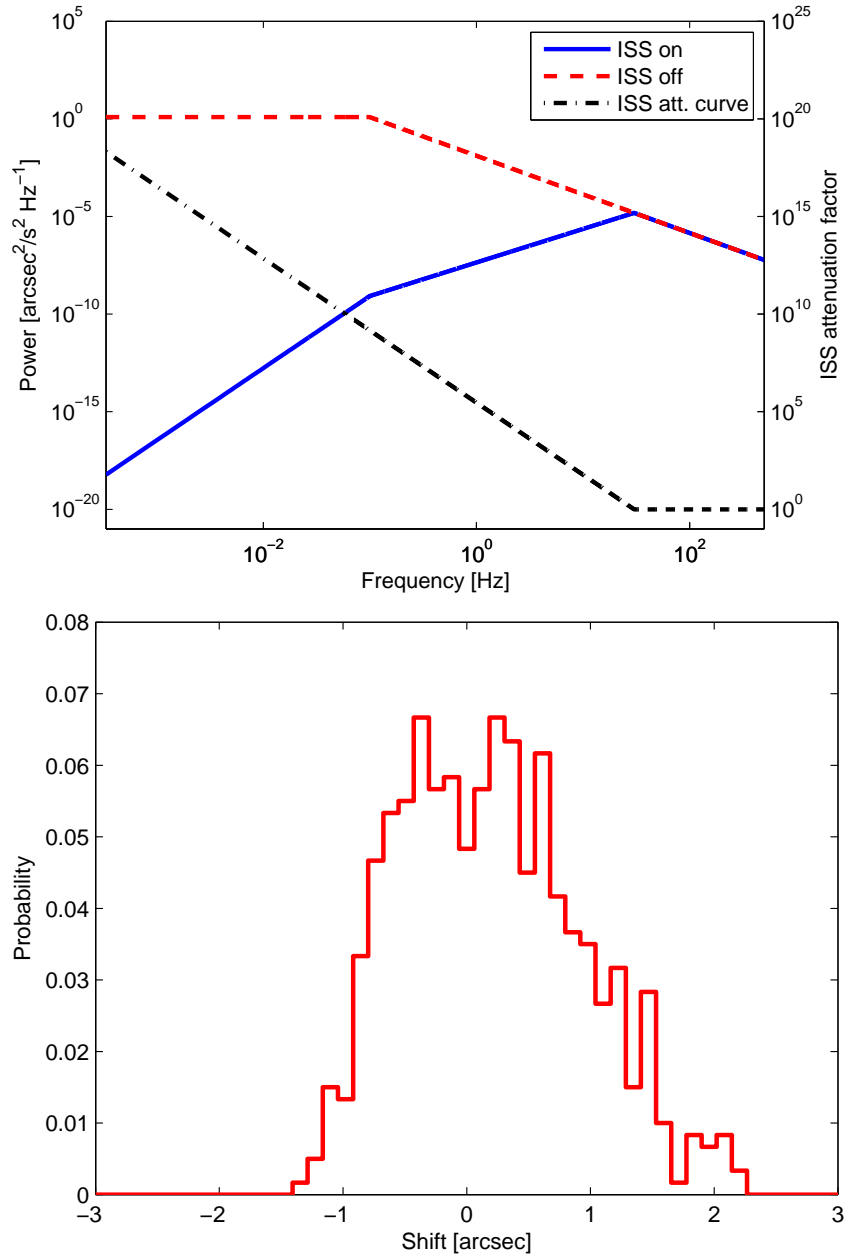


Figure 2.3: *Top*: Power spectral density of the shift introduced by the modeled spacecraft jitter. We assume the jitter to be similar to that of the *Hinode* spacecraft and to follow the spacecraft requirement (see text for details). When the ISS is turned off (*dashed curve*), we use a power law to model the power spectral density of the *Hinode* jitter at frequencies above 0.1 Hz ( $P(\nu) \propto \nu^{-1.98}$ ) and a constant power spectral density for frequencies smaller than 0.1 Hz. We model the influence of the ISS by dividing the power spectral density of the jitter (*dashed curve*) by a model for the attenuation curve of the ISS (*dash-dotted curve*). When the ISS is turned on, the jitter is reduced significantly (*solid curve*). The ISS works most efficiently at low frequencies. *Bottom*: Example of a histogram of the shifts introduced by the jitter (ISS turned off). We used a time-series with a duration of 60 s, which corresponds to the cadence of PHI. The distribution has an RMS of 0.7''.

### 2.2.6 Point Spread Function

An important characteristic of PHI is the Modulation Transfer Function (MTF). Since the MTF is mostly determined by the optical properties of the HREW when it is exposed to the environmental conditions in the orbit, the final MTF is still subject to tests and simulations. For this work, we use two different Modulation Transfer Functions corresponding to the best and the worst possible outcome. The first MTF represents a diffraction-limited instrument, we assume the Point Spread Function (PSF) to be an Airy function corresponding to the aperture of the High Resolution Telescope (14 cm). The second one is an MTF that was computed from a model of the PHI instrument, where the entrance window was located at a suboptimal position within the heat shield of the spacecraft<sup>2</sup> (see Figure 2.4). Compared to the MTF for the diffraction-limited instrument, the MTF computed from the model of PHI strongly decreases with the wavenumber and is also very asymmetric.

### 2.2.7 Photon Noise

The last step is modeling the detector. Here we resample the data to the detector plate scale (0.5'' per pixel, corresponding to  $\sim 100$  km at perihelion) and add photon noise to the data.

PHI will accumulate several exposures to generate one final image in order to increase the signal-to-noise ratio. The requirement for PHI for polarimetry is an accuracy of  $\pm 10$  G per pixel for determining the LOS magnetic field,  $\pm 200$  G for the transverse component of the magnetic field and  $\pm 15$  m/s for the LOS velocity. This demands an instrument requirement for the signal-to-noise ratio of the mean continuum of at least 1000. Here we assume the noise to follow the requirement.

Since we simulate only one image, representing the accumulation of many individual exposures, we can add the noise directly to the intensity images using a normal distribution with dispersion

$$\sigma(\lambda, x, y) = 10^{-3} \sqrt{I(\lambda, x, y) \langle I_c \rangle}. \quad (2.1)$$

Here,  $I(\lambda, x, y)$  is the intensity at a given wavelength and a given pixel and  $\langle I_c \rangle$  is the spatially averaged continuum intensity at disk center. This results in the desired signal-to-noise ratio:

$$\frac{I}{\sigma} = 10^3 \sqrt{\frac{I(\lambda, x, y)}{\langle I_c \rangle}}. \quad (2.2)$$

Since PHI will perform an onboard correction of the raw images for darks and flat-fields, we do not model their influence here. We also neglect quantization noise. The raw images provided by the detector will have a size of 14 bits per pixel, so the influence of quantization noise is far below the photon noise level. Another source of noise is caused by variations of the exposure time. The requirement for this noise is 7 m/s, demanding a target accuracy of the exposure time of 450 ppm (Appourchaux 2011).

---

<sup>2</sup>Recently, preliminary results from a study for the optimum HREW location have shown that a nearly diffraction-limited telescope under all orbit conditions is feasible.

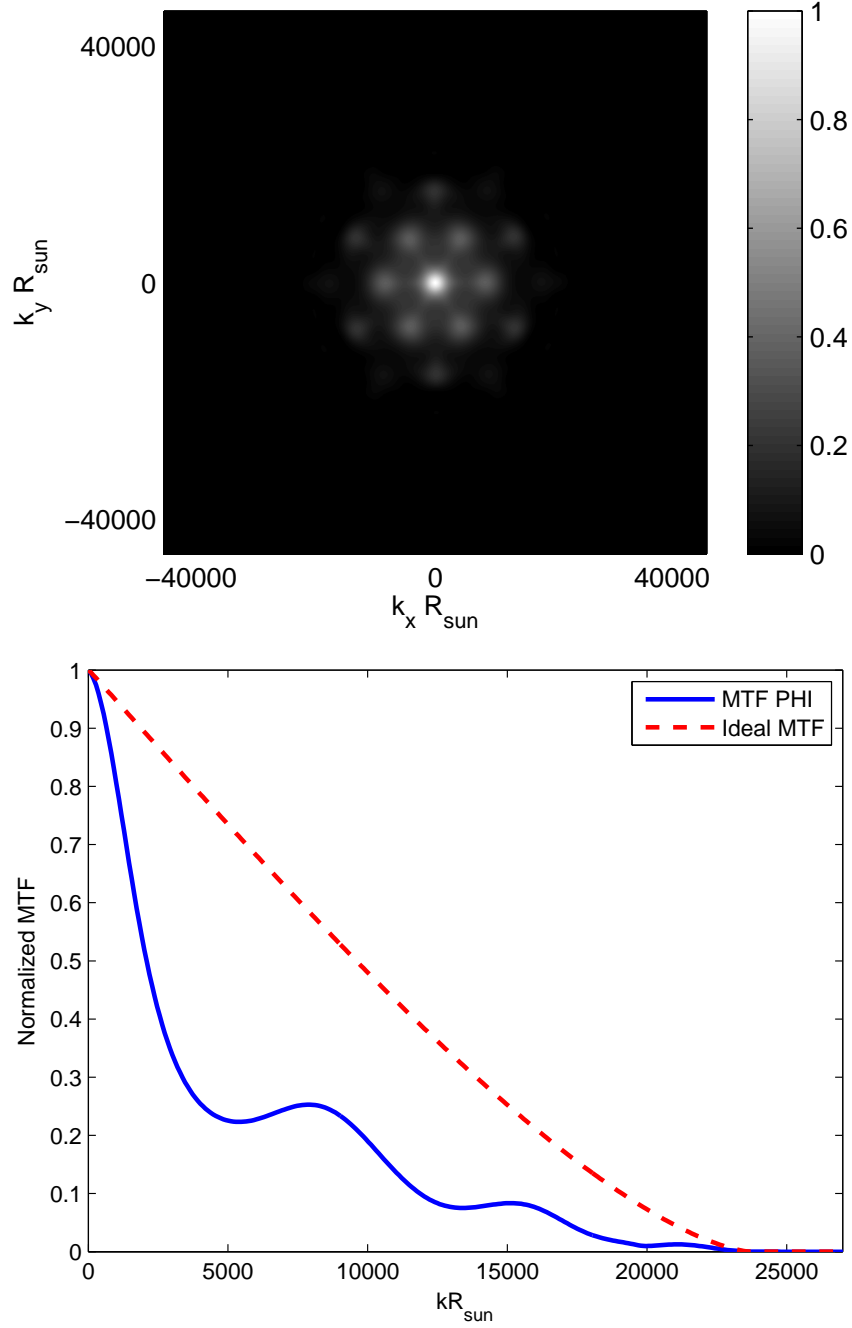


Figure 2.4: Modulation Transfer Functions used in this study as a function of wavenumber when observing at perihelion (0.28 AU). *Top*: MTF computed from a modeled instrument, where the entrance window is not designed in the best possible way, *bottom*: Azimuthal average of this MTF (*solid curve*) and an MTF corresponding to a diffraction-limited instrument (*dashed curve*).

Table 2.1: RMS for continuum intensity contrast and LOS velocity both for disk center and  $\rho = 60^\circ$ . We compare data from SPINOR with our results from SOPHISM. Both the MTF corresponding to a diffraction-limited instrument and a worst-case MTF are considered.

	RMS of intensity contrast		RMS of LOS velocity [km/s]	
	Disk center	$\rho = 60^\circ$	Disk center	$\rho = 60^\circ$
SPINOR	0.17	0.14	1.20	1.54
SOPHISM (diff.-limited MTF)	0.12	0.10	0.72	1.08
SOPHISM (worst-case MTF)	0.06	0.05	0.37	0.62

## 2.3 Synthetic Intensity and Velocity Maps

Applying the individual steps described in Section 2.2 results in a time-series of intensity images at different wavelengths as will be available for onboard processing on *Solar Orbiter*. Here we determined the LOS velocities by computing the barycenter of the intensity at the six wavelength positions that we have selected.

We generated two such time-series with a length of 359 min each, one corresponding to an observation of the disk center and another one for an observation closer to the limb (heliocentric angle  $\rho = 60^\circ$ ). This corresponds to an observation of the solar poles from a solar latitude of  $30^\circ$ , as it will be done by *Solar Orbiter*. In both cases we have assumed the spacecraft to observe from perihelion (0.28 AU). Since the length of the time-series is rather short, we have neglected spacecraft motion. Examples for intensity and velocity maps are shown in Figure 2.5 for disk center and in Figure 2.6 for  $\rho = 60^\circ$ .

At disk center, granulation is clearly visible in the intensity images and Dopplergrams. When observing at  $\rho = 60^\circ$ , the granulation can still be resolved in the intensity images but the contrast between granules and intergranular lanes is reduced. The Dopplergrams at the limb are dominated by horizontal outflows from the granules to the intergranular lanes. After running SOPHISM, the contrast of both the intensity and the velocity is reduced significantly (see Table 2.1) due to the MTF. There are strong differences between the two MTFs that we use here, the worst-case MTF reduces the RMS both for intensity and velocity by about a factor of two.

## 2.4 Oscillation Power Spectra

We have a time-series of Dopplergrams and thus we can evaluate the effect of the PHI instrument on the measurement of solar oscillations. The most basic of helioseismic diagnostics is the power spectrum as a function of horizontal wavenumber,  $k$ , and frequency,  $\nu$ . In Figure 2.7 we compare two power spectra for an observation at disk center: one generated from synthetic PHI data (i.e. after running SOPHISM) and another one without any influence of the instrument (i.e. for velocity maps derived from the SPINOR line profiles). We see that the power resulting from the synthetic PHI data is lower than the power originating from the line profiles at high- $k$  modes. This is caused by the MTF,



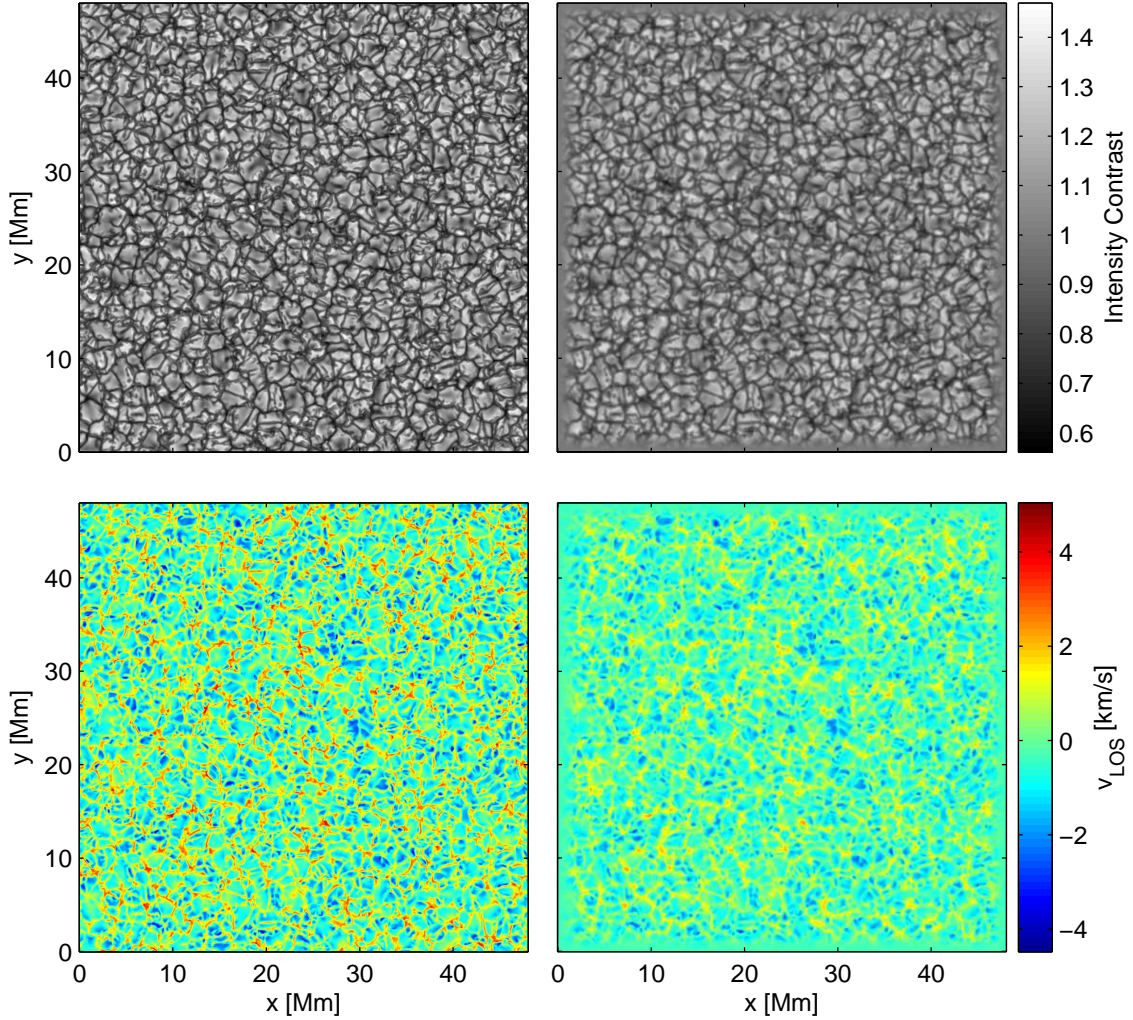


Figure 2.5: Influence of the modeled High Resolution Telescope of PHI on intensity (*top*) and line-of-sight velocity (*bottom*) when observing disk center from perihelion (0.28 AU). We compare data from the line profiles (*left*) with our results for the modeled PHI instrument (*right*). The dominant signal in these images is granulation. The Dopplergrams clearly show the upflows in the granules and the downflows in the intergranular lanes. In the modeled PHI data, the spatial resolution is reduced due to the MTF. The edges of the SOPHISM data are smooth, because SOPHISM apodizes the image before applying the MTF in order to avoid artifacts in the Fourier spectrum. Here we use an MTF corresponding to a diffraction-limited instrument. The image is only a fraction of the full FOV for the HRT.

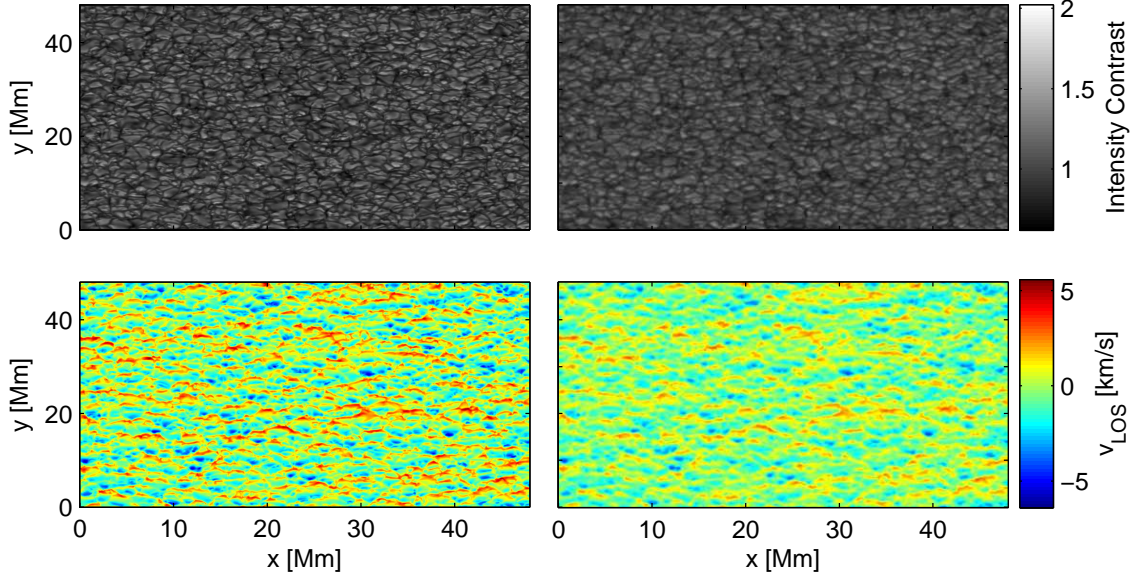


Figure 2.6: Same as Figure 2.5 for an observation at  $\rho = 60^\circ$ . Again, we assume an observation from perihelion (0.28 AU). Due to foreshortening the images are squeezed in the y-direction. The top of the box is directed towards the limb. While the intensity images show granulation, as for the disk center data, the Dopplergrams are now dominated by flows in the horizontal direction, which originate from outflows from the granules.

which we assume here to correspond to a diffraction-limited instrument.

The impact of the MTF on the power spectrum strongly differs between the two MTFs used in this study (see Section 2.2.6). Figure 2.8 shows the azimuthal averages of power spectra for different setups of the instrument. The worst-case MTF reduces the power significantly more at high wavenumbers than the diffraction-limited one. The MTF has a much larger influence on the power spectrum than other instrumental effects, like spacecraft jitter or photon noise. We also show in Figure 2.8 power spectra, where the ISS is turned off and the photon noise is increased ( $S/N = 50$ ). Photon noise and jitter only influence the power spectra if they are much worse than the PHI requirements. The photon noise leads to a constant offset in the power spectrum which is negligible compared to the power of the oscillation modes for any noise level that we expect. The spacecraft jitter reduces the power of the high- $k$  modes, if the ISS is turned off. All these effects barely affect the oscillations.

When observing at  $\rho = 60^\circ$ , the power of the p-modes is lower than at disk center, since the displacement caused by the oscillations is mostly radial. On the other hand, the power resulting from convection is increased. Foreshortening leads to a stronger influence of the MTF and spacecraft jitter. Also, the intensity is lower due to limb darkening, which leads to an increased photon noise level.

In short, the most important instrumental effect is the MTF, which attenuates p-mode power at high spatial frequencies. The two MTFs used in this study lead to strong differences in the power spectrum. By comparison, the uncertainties in the jitter or photon noise level have less of an effect. However, we find that oscillations are still easy to observe even with the worst-case MTF.



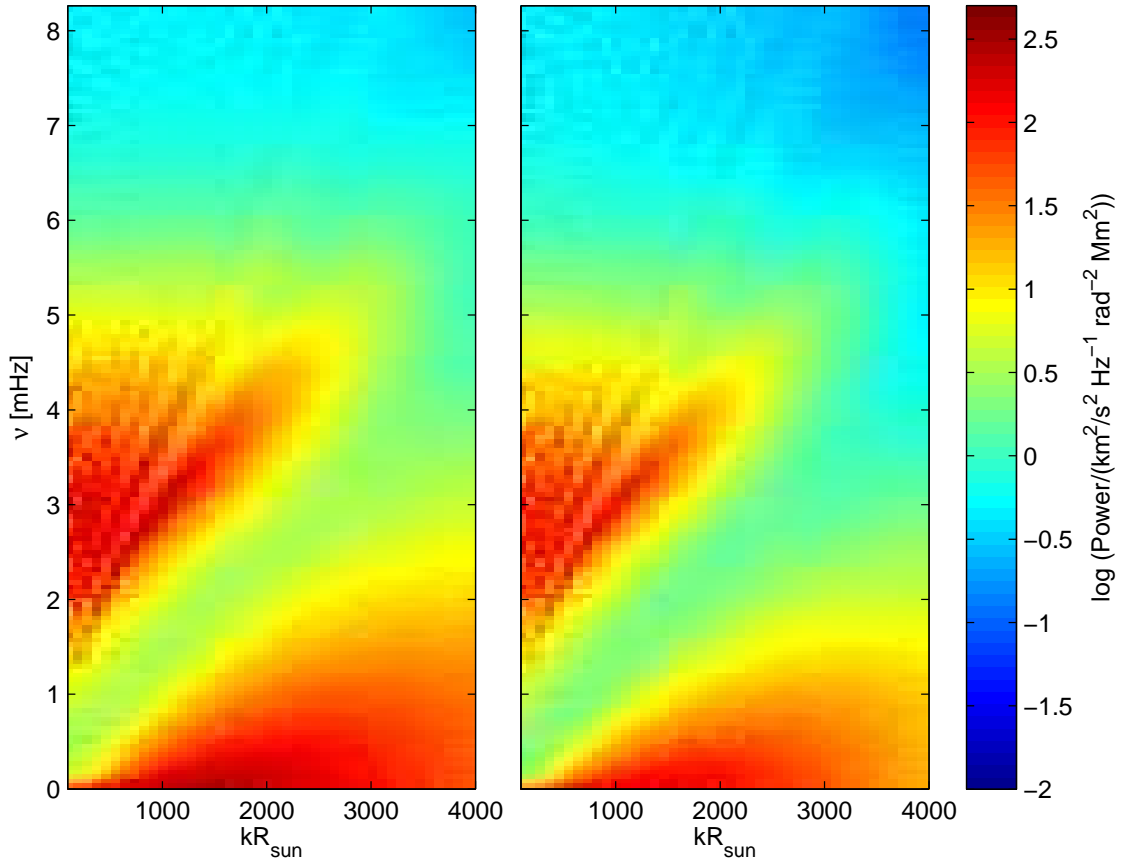


Figure 2.7: Azimuthal average of two power spectra at disk center. *Top*: Power for velocities determined directly from the line profiles, *bottom*: Power for velocities computed from simulated instrumental data after running SOPHISM. Both images reveal the solar oscillations and granulation. However, the simulated instrument decreases the power, especially for high wavenumbers. This is caused mostly by the Modulation Transfer Function. Here we use an MTF corresponding to a diffraction-limited instrument (the PSF is an Airy function).

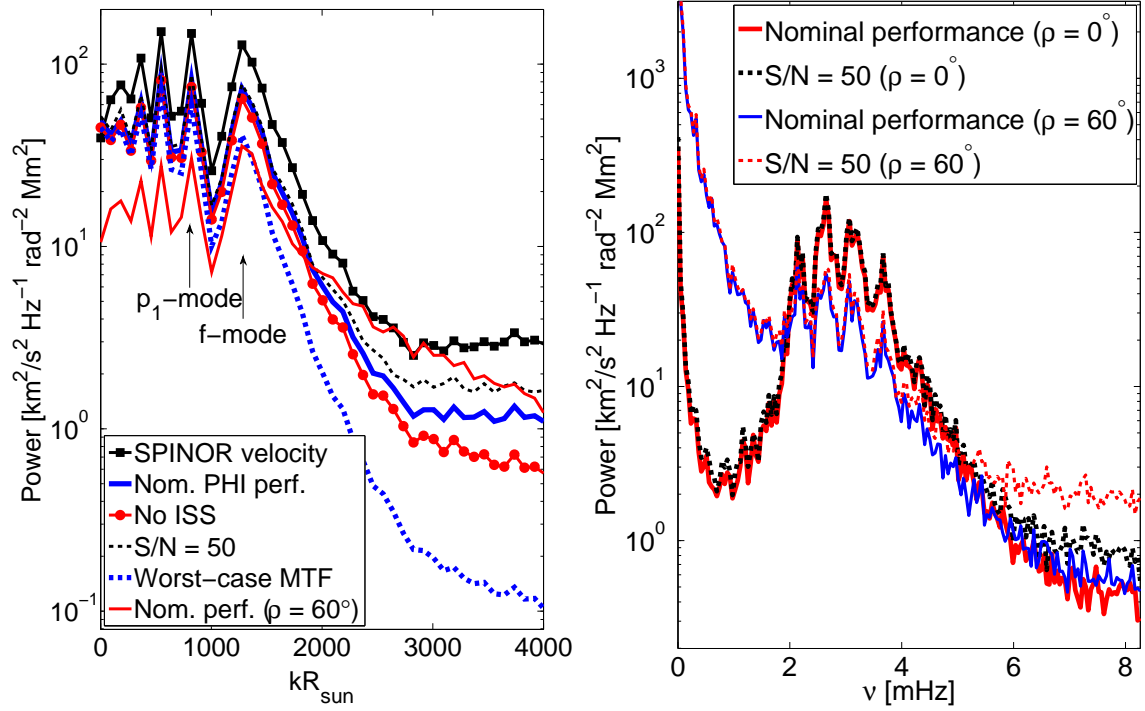


Figure 2.8: *Left*: Azimuthal average of the power as a function of wavenumber at  $\nu = 3.41$  mHz. Six configurations are shown, *black squares*: Power derived from SPINOR velocities, *blue thick solid curve*: Nominal PHI performance, *red dots*: ISS turned off, *black dashed curve*: Enhanced photon noise ( $S/N = 50$ ), *blue thick dashed curve*: Worst-case MTF, *red solid curve*: Nominal PHI performance at  $\rho = 60^\circ$ . *Right*: Power as a function of frequency at  $kR_\odot = 456$ . *Red thick solid line*: Nominal PHI performance when observing at disk center, *blue solid line*: Nominal PHI performance at  $\rho = 60^\circ$ , *black thick dashed line*: Enhanced photon noise ( $S/N = 50$ ) at disk center, *red dashed line*: Enhanced photon noise ( $S/N = 50$ ) at  $\rho = 60^\circ$ . We do not show the jitter with the ISS off, since it does not depend on frequency. The photon noise enhances the background power. For the nominal PHI performance, the influence of jitter and photon noise is negligible. If the ISS is turned off, the high- $k$  power is reduced, a higher photon noise level increases the background power. However, a signal-to-noise ratio ( $S/N$ ) of 50 is much lower than anything we expect.

## 3 Image Compression in Local Helioseismology

This chapter appeared in the paper *Image Compression in Local Helioseismology* (A&A, 571, A42, 2014, reproduced with permission from Astronomy & Astrophysics, © ESO) by B. Löptien, A. C. Birch, L. Gizon, and J. Schou. I designed and performed the research with helpful advice from the coauthors and wrote this chapter. The tracked and remapped Dopplergrams and the noise cubes were provided by J. Langfellner, the time-distance code by A.C. Birch, and the code for the Huffman compression by J. Schou.

### Chapter Summary

Several upcoming helioseismology space missions are very limited in telemetry and will have to perform extensive data compression. This requires the development of new methods of data compression. We give an overview of the influence of lossy data compression on local helioseismology. We investigate the effects of several lossy compression methods (quantization, JPEG compression, and smoothing and subsampling) on power spectra and time-distance measurements of supergranulation flows at disk center. We applied different compression methods to tracked and remapped Dopplergrams obtained by the *Helioseismic and Magnetic Imager* onboard the *Solar Dynamics Observatory*. We determined the signal-to-noise ratio of the travel times computed from the compressed data as a function of the compression efficiency. The basic helioseismic measurements that we consider are very robust to lossy data compression. Even if only the sign of the velocity is used, time-distance helioseismology is still possible. We achieve the best results by applying JPEG compression on spatially subsampled data. However, our conclusions are only valid for supergranulation flows at disk center and may not be valid for all helioseismology applications.

### 3.1 Introduction to Chapter 3

Efficient compression of helioseismic data plays an important role in space missions, where it helps reduce the amount of telemetry required for transmitting data. The currently running *Helioseismic and Magnetic Imager* (HMI, Schou et al. 2012) is not very limited in telemetry and relies on quantization and lossless compression only. On the other hand, the *Michelson Doppler Imager* (MDI) instrument (Scherrer et al. 1995) had to use extensive data compression. During the medium- $\ell$  program (Kosovichev et al. 1996,

1997), for example, the images were cropped, smoothed and subsampled, and quantized before applying lossless compression to the data.

Some upcoming missions, such as for example the *Solar Orbiter* mission (Marsch et al. 2000, Marsden & Müller 2011, Löptien et al. 2014b), will also have a low bandwidth for transferring data. These missions can, of course, benefit from the experience in compression obtained by previous missions, in particular MDI. However, compression efficiency can certainly be improved, especially when using lossy compression algorithms. The behavior of lossy methods is complex and involves a trade-off between the compression efficiency and the amount of noise caused by the compression. The impact of compression artifacts might also change for different helioseismic measurements.

Like in many other cases, helioseismology can benefit from the experience with compression in Earth seismology. When analyzing seismic noise, it is common to use the sign of the oscillations only. This helps in the analysis because it removes seismic events with large amplitudes (e.g., Aki 1965, Hanasoge & Branicki 2013).

In this work, we give a first overview of lossy data compression in local helioseismology. We test the impact of compression on two of the most basic helioseismic measurements: the power spectrum and time-distance helioseismology of supergranulation at disk center. We start from Dopplergrams obtained by the HMI instrument, apply different compression methods (quantization, JPEG compression, and smoothing and subsampling) to the data and discuss their influence on the resulting power spectra and travel time maps. This allows to give a first estimate on the quality and the efficiency of these compression methods.

## 3.2 Compression Schemes

### 3.2.1 Quantization and Huffman Encoding

Quantization compresses data by reducing the number of bits per pixel used for storing the data. The number of bits per pixel determines how many different values  $n$  the velocity can assume, e.g., for five bits per pixel  $n = 32$ . Hence, quantization reduces the precision of the velocity.

We apply quantization to the Dopplergrams by rounding the velocity to a number of possible values  $n$  which are evenly distributed between the lowest and highest velocity in the time-series. The properties of the noise introduced by the quantization depends on  $n$ . For high  $n$ , this quantization noise is spatially uncorrelated and roughly equivalent to white noise. Lower values of  $n$  lead to a dependence of the quantization noise on the input velocities, causing sharp edges in the Dopplergrams. A special case is  $n = 2$ ; this allows only to distinguish between up- and downflows relative to the mean velocity.

Additional compression can be achieved by combining quantization with a lossless compression algorithm. Quantization only decreases the number of bits per pixel that are saved to a fixed value. However, using a fixed number of bits per pixel is not very efficient, since the histogram of the velocities does not correspond to a uniform distribution. Velocities with low amplitudes are much more frequent than those with high amplitudes. We apply Huffman encoding (Huffman 1952), an entropy-based lossless compression algorithm which uses a non-constant number of bits per pixel for storing the data. More

frequent values of the velocity are saved using fewer bits per pixel than less frequent ones.

The efficiency of Huffman compression depends on the statistics of the input data. The more information there is in the data, the larger the file size of the compressed data. Since the velocities in the Dopplergrams are spatially and temporally correlated, we decrease the amount of information significantly by predicting the velocity of each pixel from the surrounding pixels both in time and space using linear regression. Standard Huffman encoding does not allow us to reach file sizes smaller than one bit per pixel because every pixel is saved separately. We avoid this by concatenating the velocities from three consecutive pixels into one symbol before using the Huffman compression.

We test quantization for different values of  $n$ , ranging from 256 to two. In reality, the velocity would probably not be rounded to a fixed number of possible values between the lowest and highest velocity in the image, but it would be truncated to a fixed precision. This is better if gradients in velocity are present across the image or, e.g., cosmic rays.

### 3.2.2 JPEG Compression

Another option for compressing data is to truncate the coefficients of some spatial transformation of the data. This leads to a loss of information about small spatial scales. A common example for such a method is JPEG (Joint Photographic Experts Group) compression (Wallace 1992).

JPEG compression divides the data into blocks with a size of  $8 \times 8$  pixels, then a discrete cosine transform (DCT) is applied to each block. The coefficients of the DCT are truncated depending on a quality factor  $q$  that has to be selected as an input parameter (between 0 and 100, with a lower factor meaning a higher compression ratio) and compressed using Huffman encoding. JPEG compression does not benefit from combining it with quantization. The sharp edges in the Dopplergrams introduced by the quantization affect the higher coefficients of the DCT and increase the file size.

We use the standard JPEG compression algorithm implemented in IDL<sup>1</sup> and test different values of the quality factor  $q$  (between 5 and 100).

### 3.2.3 Smoothing and Subsampling

Another method for reducing the size of the data is decreasing the spatial resolution by subsampling the data. Subsampling decreases the Nyquist wavenumber of the data. In our case, we use only  $2 \times 2$  subsampling on HMI data near disk center, meaning that the Nyquist wavenumber is still at  $kR_{\odot} \approx 3100$ . So, aliasing only has a small influence on the power spectrum and it is sufficient to convolve the Dopplergrams with a narrow Gaussian ( $\sigma = 0.4$  pixels) to remove artifacts resulting from aliasing. After subsampling, we apply one of the two compression methods described above to the Dopplergrams.

---

<sup>1</sup>IDL (Interactive Data Language) is a product of EXELIS Visual Information Solutions, <http://www.exelisvis.com/>

### 3.3 Results

#### 3.3.1 Input Observations

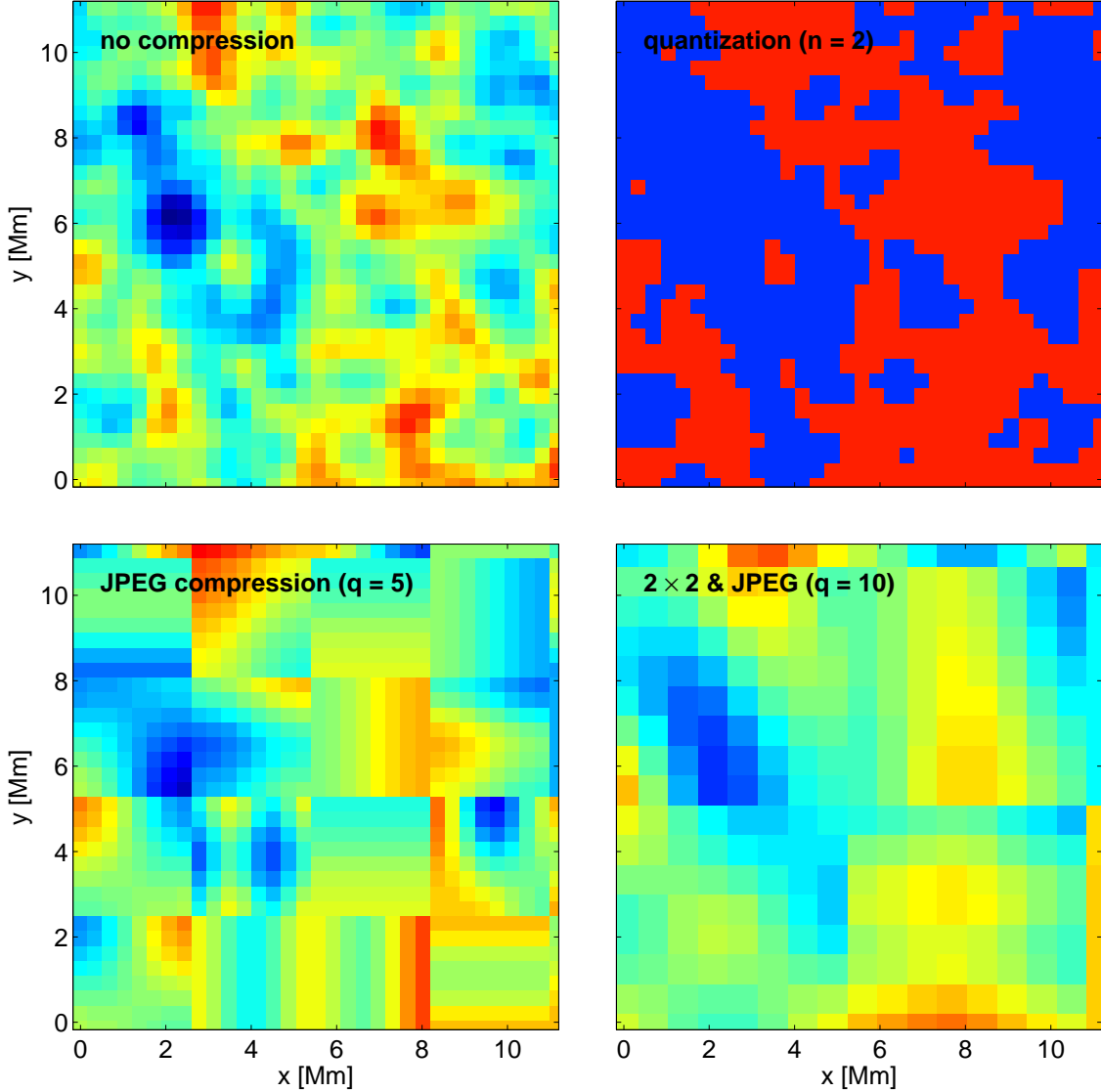


Figure 3.1: Examples of uncompressed and compressed Dopplergrams. *Top left*: uncompressed data, *top right*: quantization (two velocity bins), *bottom left*: JPEG compression (quality = 5), *bottom right*:  $2 \times 2$  subsampling combined with JPEG compression with a quality of 10. The artifacts resulting from the compression are clearly visible in the Dopplergrams. Quantization causes sharp edges in the image; if  $n = 2$ , only upflows and downflows can be separated. JPEG compression leads to blocks in the image with a size of  $8 \times 8$  pixels, corresponding to 2.78 Mm or 5.56 Mm, if  $2 \times 2$  subsampling is used. The images shown represent only a small part of the full size of the Dopplergrams ( $178 \times 178$  Mm).

Our analysis is based on Dopplergrams for the quiet Sun provided by the HMI instrument. HMI obtains filtergrams at six wavelength positions around the Fe I 6173 Å line and

transmits them to Earth, where the observables, e.g., Dopplergrams, are computed. Since HMI is basically unlimited in telemetry, the raw images that are downlinked by HMI have a size of  $\sim 7.1$  bits per pixel. Hence, the precision of the resulting Dopplergrams is far beyond the photon noise level.

We use twenty time-series of Dopplergrams obtained in May 2010 that are tracked and remapped at the equator corresponding to the Mt. Wilson 1982/84 differential rotation rate (Snodgrass 1984) around the central meridian. Each of the time-series has a length of eight hours and a size of  $178 \times 178$  Mm ( $512 \times 512$  pixels, spatial resolution: 0.348 Mm). The tracked and remapped Dopplergrams are saved in single precision (32 bits per pixel). In the next step, we apply the various compression schemes described in Section 3.2 to the data. We study a broad range of parameters (velocity bins  $n$  and quality factor  $q$ ), but the results we present in the following sections are mostly for the highest compression factors since these show the strongest influence of compression.

Examples of extremely compressed Dopplergrams are shown in Figure 3.1. They clearly show the artifacts resulting from the compression. Quantization reduces the number of possible values of the velocity, causing sharp edges in the Dopplergrams, and JPEG compression leads to artifacts in the image based on the blocks used for the JPEG compression.

### 3.3.2 Influence on Helioseismic Power Spectra

The most basic tool in helioseismology is the power spectrum as a function of wavenumber and frequency. In Figs. 3.2 and 3.3, we compare azimuthally averaged power spectra for uncompressed and examples of extremely compressed data (quantization with two bins in velocity, JPEG compression with a quality factor of five, and JPEG compression with  $q = 10$  applied to  $2 \times 2$  subsampled data).

Even for high compression factors, the p- and f-modes are still clearly visible in the power spectrum. However, depending on the compression factor, the compression decreases the power of the oscillations and adds noise. The characteristics of the compression noise depend on the method. Quantization adds an almost flat background noise to the power spectrum. The noise caused by the JPEG compression exhibits a more complicated shape. For very low quality factors, several ridges appear at high wavenumbers, depending on the spatial resolution of the data. Also, the power arising from granulation is reduced significantly.

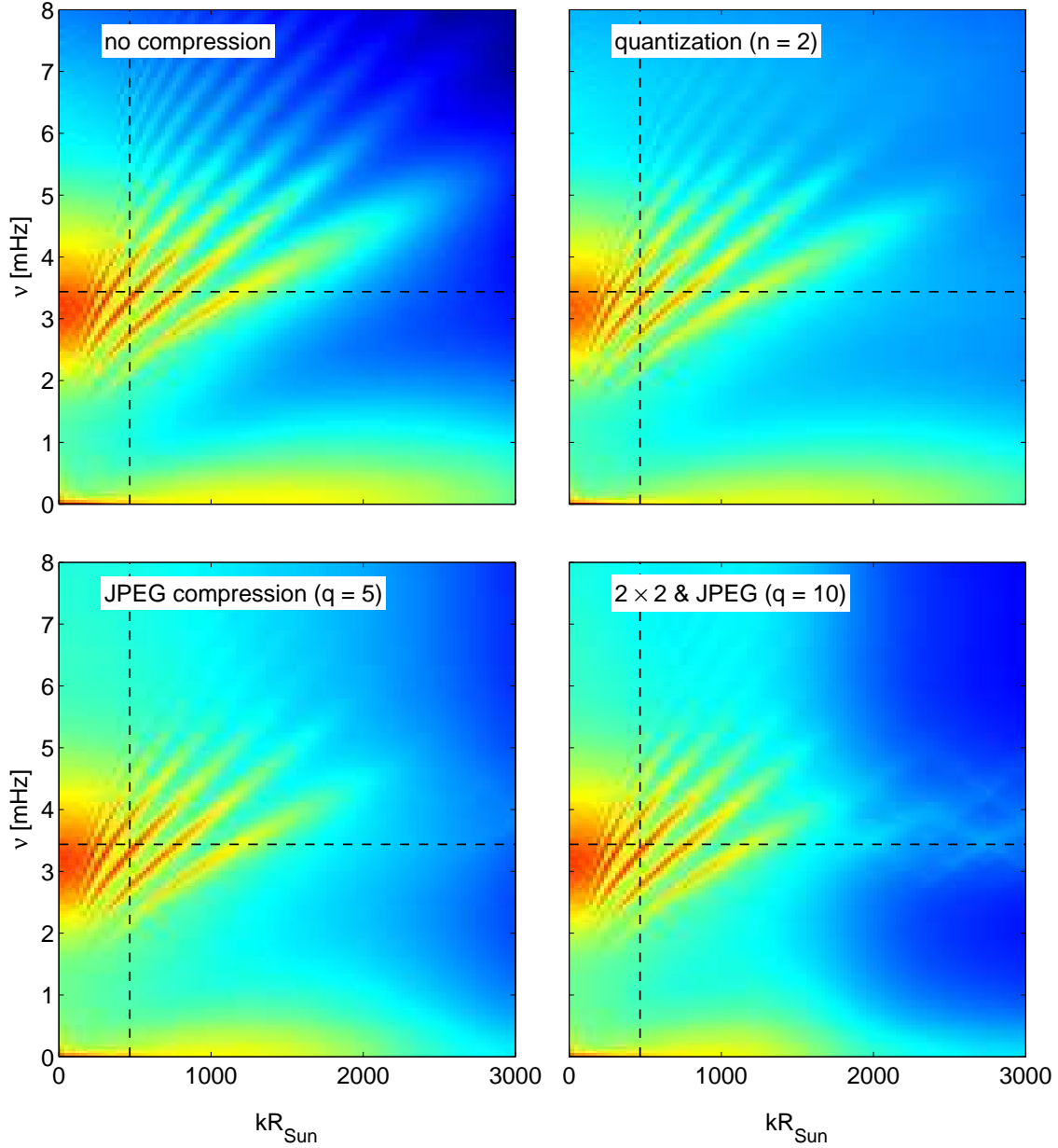


Figure 3.2: Compression artifacts in azimuthally averaged power spectra. *Top left*: Uncompressed data, *top right*: quantization (two velocity bins), *bottom left*: JPEG compression (quality = 5), *bottom right*:  $2 \times 2$  subsampling combined with JPEG compression with a quality of 10. All four power spectra are plotted using the same logarithmic color-scale (red corresponds to high power, blue to low power) and are averages from twenty time-series, each of them having a length of eight hours. We normalized the bit-truncated Dopplergrams to have the same variance as the uncompressed data before computing power spectra. All compression methods slightly decrease the power of the modes and add additional noise to the power. The noise caused by quantization is almost flat, the noise caused by JPEG compression exhibits a more complex behavior, including several ridges appearing at high wavenumbers and a reduction of the power arising from granulation. In Figure 3.3, we show cuts through the power spectra along the dashed lines.



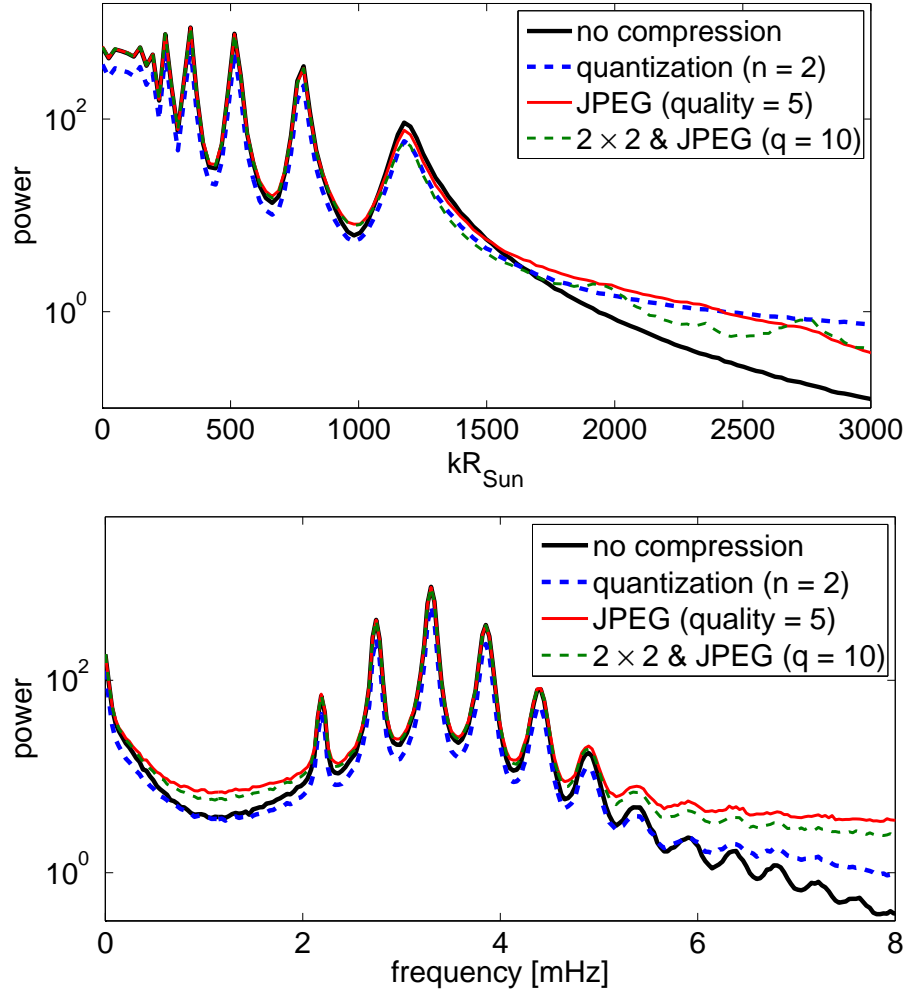


Figure 3.3: Cuts through the azimuthally averaged power spectra shown in Figure 3.2 (cut along the dashed lines). *Top*: power at a constant frequency,  $\nu = 3.4$  mHz, *bottom*: power at a constant wavenumber,  $kR_{\odot} = 466$ . Four configurations are shown. *Thick solid black curve*: no compression is applied to the data, *thick dashed blue curve*: the Dopplergrams are compressed using quantization (two velocity bins), *solid red curve*: JPEG compression (quality = 5), *dashed green curve*:  $2 \times 2$  subsampling and JPEG compression with a quality of 10. The power shown is an average of twenty time-series, each of them having a length of eight hours. We normalized the bit-truncated Dopplergrams to have the same variance as the uncompressed data before computing power spectra. All compression methods slightly decrease the power of the modes and add additional noise to the power.

### 3.3.3 Influence on Supergranulation Travel Times

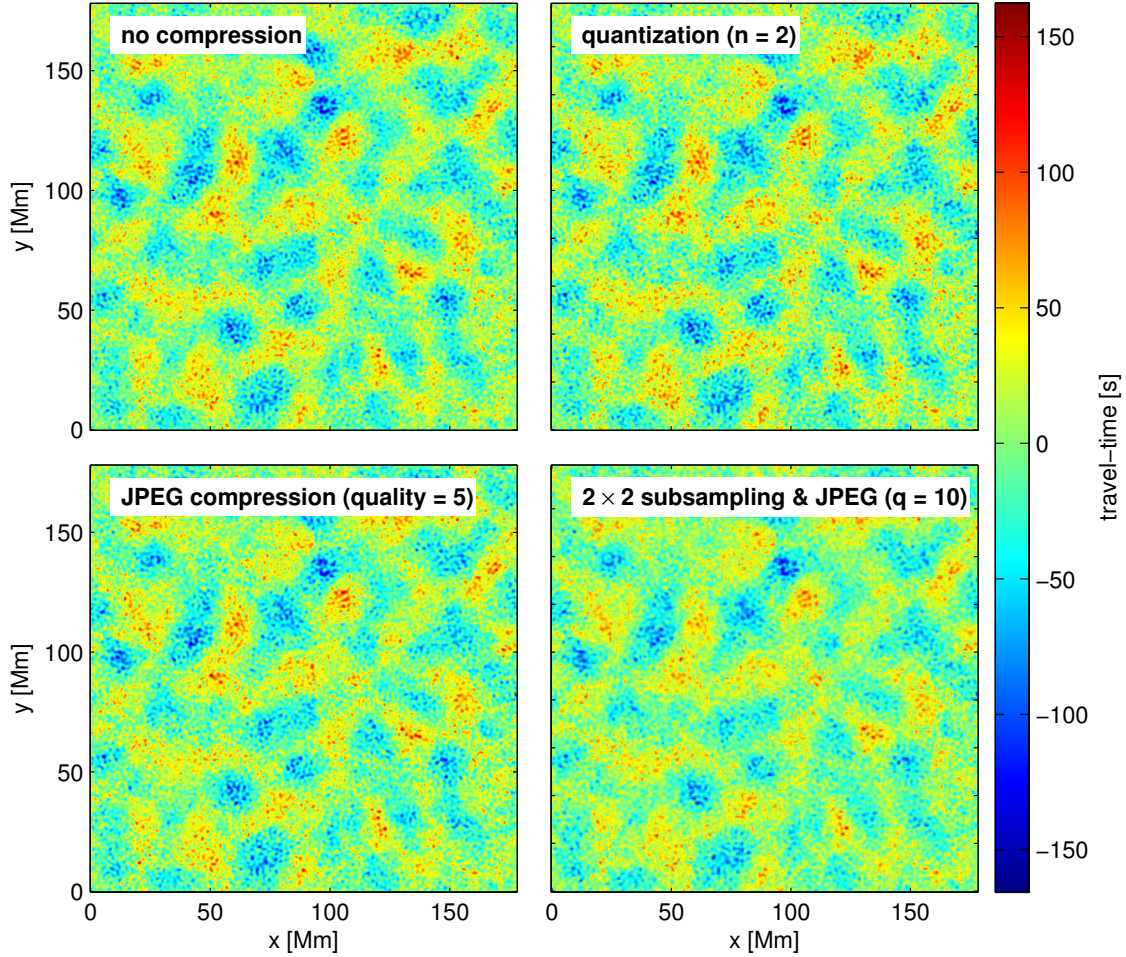


Figure 3.4: Examples of travel time maps of supergranulation (f-mode, center-to-annulus geometry, annulus size 13.4 Mm, outward minus inward travel times) computed from compressed and uncompressed Dopplergrams. *Top left*: Uncompressed data, *top right*: quantization (two velocity bins), *bottom left*: JPEG compression (quality = 5), *bottom right*:  $2 \times 2$  subsampling combined with JPEG compression with a quality of 10. The travel times clearly show the supergranulation pattern. They correspond to the divergence of horizontal flows, with negative values indicating outflows and positive values showing inflows.

We evaluate the influence of the compression methods described in Section 3.2 on time-distance helioseismology (Duvall et al. 1993) by computing travel time maps for the f-mode using point-to-annulus geometry (annulus size 13.2 Mm) and by taking the difference in travel times between outgoing and ingoing waves. These travel times are sensitive to the horizontal divergence of near-surface flows (Duvall & Gizon 2000), such as supergranulation. Examples of the resulting travel time maps are shown in Figure 3.4.

Although the compression introduces additional noise, the travel time maps for quantization and JPEG compression are almost indistinguishable from the uncompressed case. The correlation between uncompressed and compressed travel times is only for extreme

compression ratios smaller than 0.99 (see Table 3.1). There is also no apparent correlation of the noise caused by compression with the travel times. When deciding on a compression method, two things are important: the quality of the compressed data and the compression efficiency. These depend on the parameters of the compression (the number of velocity bins  $n$  of the quantization or the quality factor  $q$  of the JPEG compression). Determining the compression efficiency is straightforward; here we determine the quality of the compressed data by comparing the signal-to-noise ratio (S/N) with the uncompressed data.

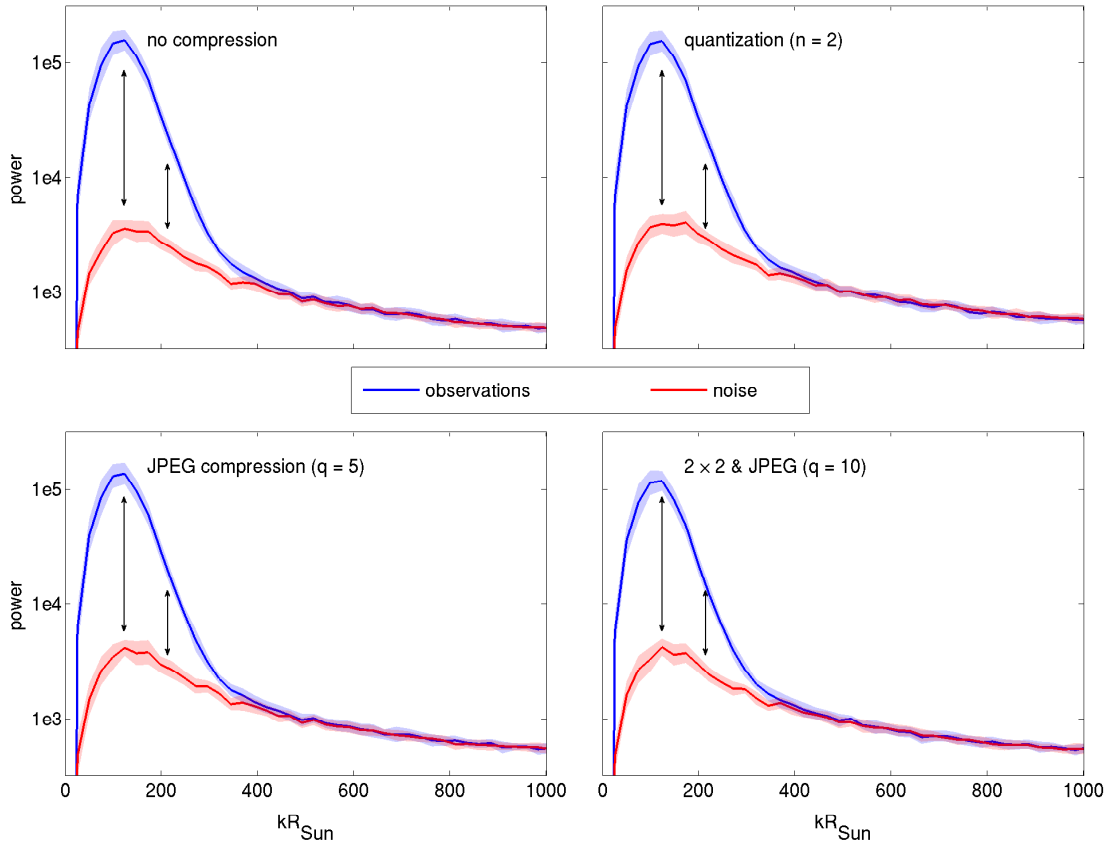


Figure 3.5: Azimuthally averaged power of travel times (center-to-annulus geometry, outward minus inward travel times) derived from observations (*blue curve*) and a noise model (*red curve*). *Top left*: Uncompressed data, *top right*: quantization (two velocity bins), *bottom left*: JPEG compression (quality = 5), *bottom right*:  $2 \times 2$  subsampling combined with JPEG compression with a quality of 10. The power shown is an average from twenty time-series, each of them having a length of eight hours. The shaded areas show the  $1\sigma$  scatter of the individual realizations. The compression slightly decreases the signal and increases the noise. The vertical arrows denote the wavenumbers used for determining the S/N in Figure 3.6 and Table 3.1.

The S/N of travel time maps is defined as the difference of the power of the travel times computed from the observations and a noise model divided by the noise (Gizon & Birch 2004, Fournier et al. 2014).

Both the observations and the noise model are affected by compression. As can be

Table 3.1: Properties of travel time maps (center-to-annulus geometry, outward minus inward travel times) computed from compressed data. We evaluate the quality of the compression for selected compression efficiencies (number of velocity bins  $n$  or quality factor  $q$ ) by computing the correlation of the travel times with the travel times computed from uncompressed data, the signal-to-noise ratio (S/N) of the travel times for two wavenumbers ( $kR_{\odot} = 122$  and  $kR_{\odot} = 221$ ) and the file size of the Dopplergrams in bits/pixel (relative to the data with full spatial resolution) after applying the compression. The  $2 \times 2$  subsampled data were interpolated back to the original resolution using Fourier interpolation for computing the correlations. Before determining the S/N, we averaged the power of the travel times from  $kR_{\odot} = 98 - 147$  and  $kR_{\odot} = 196 - 245$ , respectively. Our measurement of the S/N for the uncompressed case has an uncertainty of about 2. The other error estimates are for differential S/N measurements relative to the uncompressed case.

compression method	parameter	Corr.	S/N at $kR_{\odot} = 122$	S/N at $kR_{\odot} = 221$	file size [bits/pixel]
no compression quantization	$n = 2$	1	$40 \pm 2$	$8.0 \pm 0.3$	32
	$n = 4$	0.965	$35.1 \pm 0.8$	$6.5 \pm 0.1$	0.71
	$n = 8$	0.975	$36.8 \pm 0.7$	$6.81 \pm 0.09$	0.72
JPEG compression	$n = 8$	0.998	$39.8 \pm 0.2$	$7.84 \pm 0.03$	0.92
	$q = 5$	0.957	$31 \pm 1$	$5.8 \pm 0.2$	0.23
	$q = 10$	0.991	$38.2 \pm 0.4$	$7.50 \pm 0.08$	0.43
$2 \times 2$ subsampling and quantization	$q = 20$	0.998	$39.2 \pm 0.3$	$7.91 \pm 0.03$	0.71
	$n = 2$	0.948	$32 \pm 1$	$5.9 \pm 0.2$	0.23
	$n = 4$	0.960	$34 \pm 1$	$6.2 \pm 0.2$	0.24
$2 \times 2$ subsampling and JPEG compression	$n = 8$	0.989	$37.9 \pm 0.6$	$7.36 \pm 0.09$	0.30
	$q = 5$	0.612	$7 \pm 2$	$0.8 \pm 0.3$	0.06
	$q = 10$	0.927	$27 \pm 2$	$4.6 \pm 0.2$	0.12
	$q = 15$	0.972	$34.2 \pm 0.9$	$6.0 \pm 0.2$	0.17

seen in Figure 3.5, the compression slightly decreases the power caused by the supergranulation and enhances the effect of the realization noise. This decreases the S/N.

We determine the S/N at two wavenumbers which correspond to the length scale of supergranulation ( $kR_{\odot} = 122$  and  $kR_{\odot} = 221$ , indicated by the vertical arrows in Figure 3.5). In Table 3.1 and in Figure 3.6, we show the resulting S/N and the file sizes for the different compression methods while varying the parameters of the compression (the number of velocity bins  $n$  for the quantization and the quality factor  $q$  of the JPEG compression). Since the behavior of the compression methods is very similar at these two wavenumbers, we show only the S/N at  $kR_{\odot} = 122$  in the figure.

All compression methods decrease the file size significantly; a compression to a file size of one bit per pixel is possible without any major influence on the S/N. However, there are large differences between the individual methods. JPEG achieves, in general, a better S/N than the combination quantization and Huffman encoding, especially for high compression factors. A quality factor of ten leads to a S/N of 38 at  $kR_{\odot} = 122$  which is comparable to the S/N of the uncompressed data ( $S/N = 40$ ). This corresponds to a file size of 0.43 bits per pixel. Additional compression can be achieved by using JPEG compression on  $2 \times 2$  subsampled data. This is the best method tested in this study. No matter which S/N is required (except for a  $S/N > 39.5$ ), this compression scheme leads to the smallest file size. A further reduction of the spatial resolution is not advisable, using  $4 \times 4$  subsampling significantly reduces the S/N (the maximum S/N is 20). The spatial resolution of  $4 \times 4$  subsampled data is  $4 \times 0.348 = 1.392$  Mm. This is comparable to the typical wavelength of the f-mode ( $\sim 4 - 8$  Mm);  $4 \times 4$  subsampling might work better for p-modes which have at the same frequency a larger wavelength than the f-mode.

### 3.4 Discussion and Conclusion for Chapter 3

Probing supergranulation flows at disk center seems to be very robust to data compression. However, the good performance of the compression is probably not valid for helioseismology in general. There are helioseismic measurements with a much lower S/N than the simple measurements presented in this paper, e.g., weak and deep flows and helioseismology at high latitudes. Most likely, these are much more sensitive to data compression. Also, center-to-limb effects could have a strong effect here. Subsampling and JPEG compression should presumably both be affected by foreshortening when applied close to the limb. The effect of compression on other measurement techniques also needs to be studied. For example, the reduction of the power arising from granulation by the JPEG compression indicates that granulation tracking will be affected by this compression method. Hence, more work is needed in order to derive a strategy for data compression that can be applied to all sorts of helioseismic measurements.

While the excellent performance of the compression for travel time measurements might be surprising at first, it can be explained by looking at the power spectra. Most of the compression artifacts are located at higher wavenumbers than used in time-distance helioseismology. This means that the filtering applied for the travel time measurements removes most of the compression noise. The noise at the location of the modes is small compared to the power of the oscillations. Hence, it has only minor influence on travel time measurements.

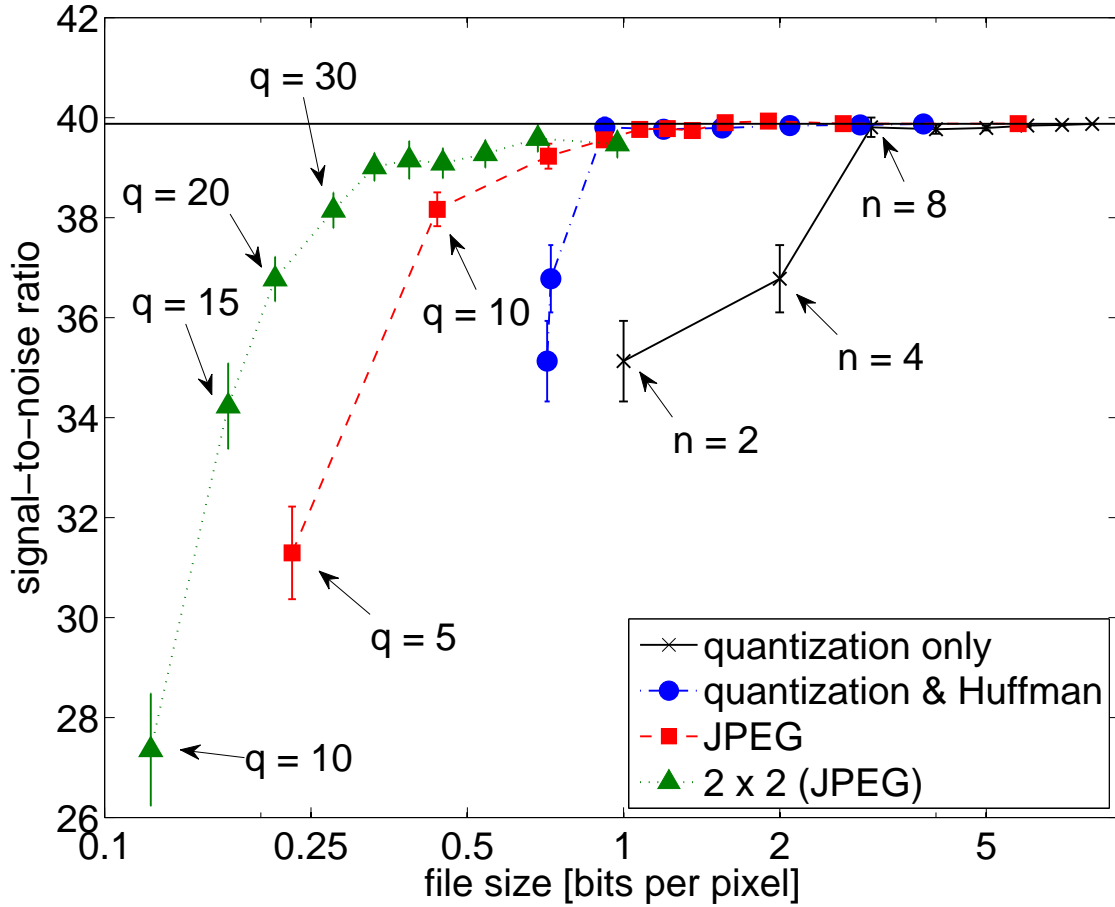


Figure 3.6: S/N of travel times (center-to-annulus geometry, outward minus inward travel times) at a wavenumber of  $kR_{\odot} = 98 - 147$  as a function of the file size in bits per pixel (relative to the data with full spatial resolution). We show data for quantization and JPEG compression using both the full spatial resolution and  $2 \times 2$  subsampled data. *Black crosses*: quantization only (full resolution), *blue circles*: quantization and Huffman compression (full resolution), *red squares*: JPEG compression (full resolution), *green triangles*: JPEG compression and  $2 \times 2$  subsampling. We vary the number of possible values for the velocity  $n$  for the quantization and the quality factor  $q$  of the JPEG compression (as indicated by the arrows). We do not show the S/N for JPEG compression with  $q = 5$  for JPEG applied to  $2 \times 2$  subsampled data because it is extremely low (S/N = 7). The S/N shown here is an average computed from twenty realizations. The horizontal line shows the S/N of the uncompressed data. Before computing the S/N, we averaged the power of the travel times from  $kR_{\odot} = 98 - 147$ . Our measurement of the S/N for the uncompressed case has an uncertainty of about 2. The error bars show error estimates for differential S/N measurements relative to the uncompressed case.

Based on the results shown in Figure 3.6, subsampling in combination with JPEG compression seems to be the best of the methods tested here for compressing data for local helioseismology. These results are, of course, limited to HMI. Helioseismic analyses require a minimum spatial resolution, so there is an upper limit on the amount by which the data can be subsampled, depending on the spatial resolution of the instrument and the

wavelength of the target waves. For HMI,  $2 \times 2$  subsampling seems to be a good trade-off between file size and image quality. In combination with JPEG compression (with varying quality factor), this reduces the file size significantly. Which compression factor can be achieved depends on the science goal of the analysis. If a low S/N is sufficient (e.g., for a statistical analysis of supergranulation), the data can be compressed down to  $\sim 0.15$  bits per pixel (relative to the full spatial resolution). Even if a high S/N is required, a compression to  $\sim 0.3$  bits per pixel is possible without decreasing the quality of the data too much.

The file sizes can probably be decreased even further. The compression methods presented in this study are very simple and probably not optimal. JPEG compression, for example, is designed to work for photos, not for scientific applications. The efficiency of JPEG could probably also be increased by not only using the spatial dimensions but also including the time-domain in the transformation. This also increased the efficiency of the Huffman compression.





# 4 Data Compression for Local Correlation Tracking of Solar Granulation

This chapter has been submitted to A&A as *Data Compression for Local Correlation Tracking of Solar Granulation* by B. Löptien, A. C. Birch, T. L. Duvall Jr., L. Gizon, and J. Schou. I designed and performed the research with helpful advice from the coauthors and wrote the chapter. The code for the Huffman compression was provided by J. Schou and the tracked and remapped HMI continuum intensity images by J. Langfellner. K. Nagashima provided comparative data on differential rotation.

## Chapter Summary

Several upcoming and proposed space missions, such as *Solar Orbiter*, will be limited in telemetry and thus require data compression. We test the impact of data compression on local correlation tracking (LCT) of time-series of continuum intensity images. We evaluate the effect of several lossy compression methods (quantization, JPEG compression, and a reduced number of continuum images) on measurements of solar differential rotation with LCT. We apply the different compression methods to tracked and remapped continuum intensity maps obtained by the *Helioseismic and Magnetic Imager* (HMI) on-board the *Solar Dynamics Observatory*. We derive 2D vector velocities using a local correlation tracking code and determine the additional bias and noise introduced by compression to differential rotation. We find that probing differential rotation with LCT is very robust to lossy data compression when using quantization. Our results are severely affected by systematic errors of the LCT method and the HMI instrument. The sensitivity of LCT to systematic errors is a concern for *Solar Orbiter*.

## 4.1 Introduction to Chapter 4

Local correlation tracking of granulation (LCT, November & Simon 1988) is an important method for measuring flows in the photosphere such as supergranulation or large-scale flows. It will play a significant role in upcoming and proposed space missions, such as *Solar Orbiter* (Marsch et al. 2000, Marsden & Müller 2011, Löptien et al. 2014b). Unfortunately, the data rate of *Solar Orbiter* will be limited and data compression will be required.

In this paper, we evaluate how continuum intensity images can be compressed with as little influence on the derived LCT flow maps as possible. We use data provided by the *Helioseismic and Magnetic Imager* (HMI, Schou et al. 2012) and the local correlation tracking code FLCT (Welsch et al. 2004, Fisher & Welsch 2008). We focus on measuring differential rotation, one of the key science objectives of *Solar Orbiter*. We combine two different approaches for compressing the data. In the first case, the individual continuum images are compressed using quantization or JPEG compression, similar to what we tried in time-distance helioseismology (Löptien et al. 2014a). In the second case, the time lag between the intensity images used by the LCT code is increased, reducing the number of images. These two approaches allow us to obtain first estimates of the performance of these compression methods.

## 4.2 Data and Methods

### 4.2.1 Input Data

The analysis is based on 120 days of tracked and remapped HMI continuum intensity images from 1 May to 28 August 2010 (provided by Langfellner et al. 2015). The cubes have a size of  $178 \times 178$  Mm ( $512 \times 512$  pixels with 0.348 Mm pixel size) and are centered around six latitudes along the central meridian ( $0^\circ, 10^\circ, 20^\circ, 40^\circ, 60^\circ$  north and  $60^\circ$  south) and at  $60^\circ$  longitude east and west at the equator. The cubes have a cadence of 45 s and are tracked for 24 h each using the Mt. Wilson 1982/84 differential rotation rate (Snodgrass 1984) corresponding to the central latitude of each cube.

### 4.2.2 The FLCT Code

The FLCT code (Fourier Local Correlation Tracking, Welsch et al. 2004, Fisher & Welsch 2008) computes 2D maps of horizontal flows by tracking the motion of small features on the Sun, in our case granulation in HMI continuum images. The FLCT code has mostly been used for studying the evolution of magnetic fields in the photosphere (e.g., Kumar et al. 2010, Orozco Suárez et al. 2012).

The code measures flows by cross-correlating pairs of continuum intensity images separated in time. First, the two images are split into subimages which are apodized with a Gaussian with a width  $\sigma$ . Then, the 2D cross-covariances in  $x$  (east-west direction, increasing westwards) and  $y$  (north-south direction, increasing northwards) between the subimages are computed. The position of the maximum of the cross-covariance function is the shift between the two subimages and can be used to determine the flow velocities. The parameter  $\sigma$  defines the resolution of the resulting flow maps. Here we use  $\sigma = 6$  pixels ( $\sim 2.1$  Mm). The output of the FLCT code is a flow map for every pair of consecutive images in the input data. Running FLCT on the data described in Section 4.2.1 results in a 24 h long time-series of LCT flow maps with a cadence of 45 s. In the following, we use averages of the flow maps over 24 h.

### 4.2.3 Compression Methods

#### 4.2.3.1 Quantization

Quantization compresses the data by dividing them by a scaling factor and rounding to the nearest integer. This reduces the number  $n$  of possible values that the intensity can assume and thus the number of bits per pixel needed for storing the data. The lower  $n$ , the lower the precision of the data and the smaller the file size.

We apply quantization to the data using fixed values of  $n$  (between 2 and 256) by varying the scaling factor. Before applying the quantization, we subtract a mean image from the data in order to remove limb darkening. Limb darkening has the largest contribution to the intensity contrast at high latitudes. Without removing the mean image, quantization would predominantly sample limb darkening but not granulation.

As described in Löptien et al. (2014a), we reduce the file size further by running lossless Huffman encoding (Huffman 1952) on the quantized data. Surrounding pixels in space and time are highly correlated. We use this to increase the efficiency of the compression.

#### 4.2.3.2 JPEG Compression

JPEG (Joint Photographic Experts Group) compression (Wallace 1992) reduces the file size of images by removing information about small spatial scales. JPEG compression divides each continuum image into  $8 \times 8$  pixel blocks and applies a discrete cosine transform (DCT) to every block. The coefficients of the transformation are then truncated depending on a quality factor  $q$  (between 0 and 100, with a lower factor causing a higher compression). After truncating, the coefficients of the DCT are compressed using lossless Huffman compression.

Like for quantization, we subtract the mean image before applying JPEG compression to the data in order to remove limb darkening. We test different values of the quality factor  $q$  (between 10 and 100).

#### 4.2.3.3 Sampling Rate

Obtaining LCT flow maps every 45 s is not very efficient in reducing the noise level. The main source of noise in LCT arises from the proper motion of granules. Granules have a lifetime of about ten minutes and so, their proper motion in LCT flow maps that are separated by 45 s is highly correlated. Increasing the time lag  $\Delta T$  between consecutive flow maps only slightly increases the noise in the final flow map averaged over 24 h, as long as  $\Delta T$  is smaller than the granulation lifetime. Long time lags, however, increase the granulation noise in the final time-averaged flow map significantly.

This method does not compress the individual intensity maps but reduces the number of intensity maps required for running the LCT code. Instead of using all the images provided by HMI (separated by 45 s), we use pairs of images with the images within the pairs being separated by 45 s, but consecutive pairs by a longer time  $\Delta T$  (between 45 s and 180 min). We also test the combination of this method with compressing the individual intensity images by using quantization or JPEG compression.

The time lag  $\Delta T$  slightly decreases the efficiency of the Huffman compression, because consecutive intensity images are not highly correlated anymore.

#### 4.2.4 Calibration of LCT Velocities

In general, LCT tends to underestimate flow velocities (see e.g., Švanda et al. 2007, Verma et al. 2013). For measuring rotation, the amplitudes of the velocities are important. These are corrected by generating calibration curves for the LCT velocities. Our calibration data consist of HMI continuum intensity data, to which we add a constant flow in the  $x$ -direction by shifting the individual images using Fourier interpolation. Most the effects that could potentially cause the low amplitudes of the LCT velocities are also present in the calibration data. The flows determined from these data with LCT are a superposition of the known input velocity and the (not known) solar flows (supergranulation, rotation). In order to determine the contribution of the calibration signal to the LCT velocities, a calibration curve must be generated by shifting the same datacube using different input velocities. The contribution from solar flows should be the same for all calibration velocities. The slope of the LCT velocities as a function of the input velocities can then be used to correct for the underestimation of velocities by LCT. Example calibration curves are shown in Figure 4.1. The slope of the curve depends slightly on the distance to disk center and is severely affected by compression. Quantization and JPEG reduce the sensitivity of LCT, a time lag  $\Delta T$  does not change the slope. This calibration method assumes that the sensitivity of LCT to the actual solar flows does not depend on the spatial scale.

### 4.3 Results

#### 4.3.1 Individual Flow Maps

Figures 4.2 to 4.4 show flow velocities in the  $x$ -direction determined by local correlation tracking, both for uncompressed data and for the various compression methods discussed in Section 4.2.3. In all flow maps, supergranulation is clearly visible, but the noise level is much higher for the compressed data. Quantization and JPEG compression cause noise in the cross-covariances computed by the FLCT code, leading to additional noise in the derived velocities on a pixel-by-pixel scale. When a large time lag  $\Delta T$  between pairs of intensity images is used, the noise from the proper motion of the granules remains in the flow maps. Quantization and JPEG compression also reduce the amplitude of the velocity, but this can be compensated using the calibration curves discussed in Section 4.2.4. The increased noise level is also visible in spatial power spectra of the flow divergence (Figure 4.5), where it increases the power at high wavenumbers. However, at wavenumbers corresponding to the length scale of supergranules ( $kR_{\odot} \approx 120$ ), the influence of quantization and JPEG compression is negligible. The performance of JPEG compression is worse than quantization. It increases the noise level more than quantization and is also less correlated with the uncompressed case.

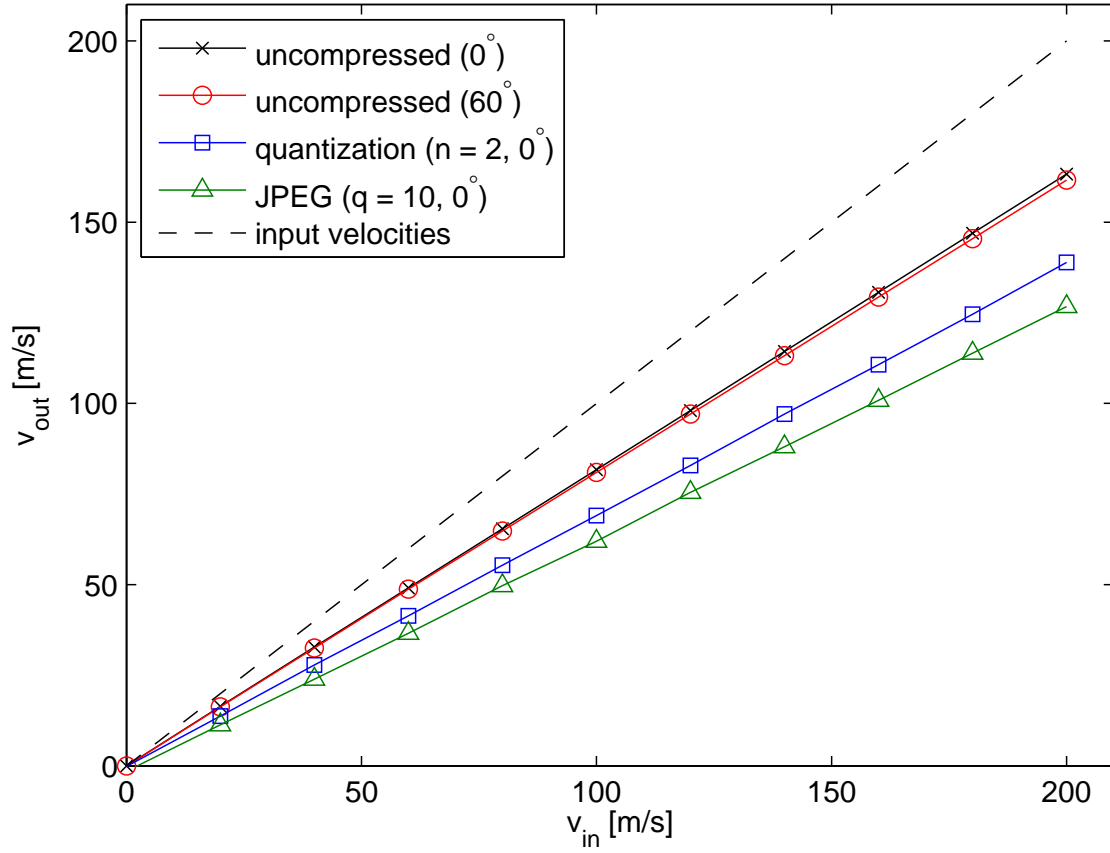


Figure 4.1: Calibration curves for LCT, *black crosses*: uncompressed data at  $0^\circ$  latitude, *red circles*: uncompressed data at  $60^\circ$  latitude, *blue squares*: quantized data ( $n = 2$ ) at  $0^\circ$  latitude, *green triangles*: JPEG compression ( $q = 2$ ) at  $0^\circ$  latitude. The error bars are smaller than the symbol size. The amplitudes of the velocities generated by LCT are in general lower than the actual velocities on the Sun, especially for compressed data. We correct for this effect by generating calibration data with a constant flow in the  $x$ -direction (as shown by the *dashed curve*). Three days of data are used for each data point. For more details about this method see the text.

### 4.3.2 Shrinking-Sun Effect

Flow maps derived by LCT are severely affected by the shrinking-Sun effect, an artifact in LCT velocities that looks like a flow pointing towards disk center (Lisle & Toomre 2004). As can be seen in Figure 4.6, it dominates the velocities derived from LCT. The amplitude of the velocities caused by this effect increases with increasing distance from disk center. This “flow” reaches up to 1.0 km/s. The relative contribution of this effect to flows in the  $x$ - and the  $y$ -direction depends on the position on the disk. Along the central meridian, for example, the main component is in the  $y$ -direction. Close to the eastern and western sides of the tracked regions, the flow component in the  $x$ -direction also becomes significant. At  $60^\circ$  latitude north, this leads to a variation of  $v_x$  of about 80 m/s in the east-west direction. In theory, the shrinking-Sun effect should not affect measurements of rotation: it should be antisymmetric around the central meridian and cancel out when averaging over longitude. However, due to the large amplitude of this effect, even a small

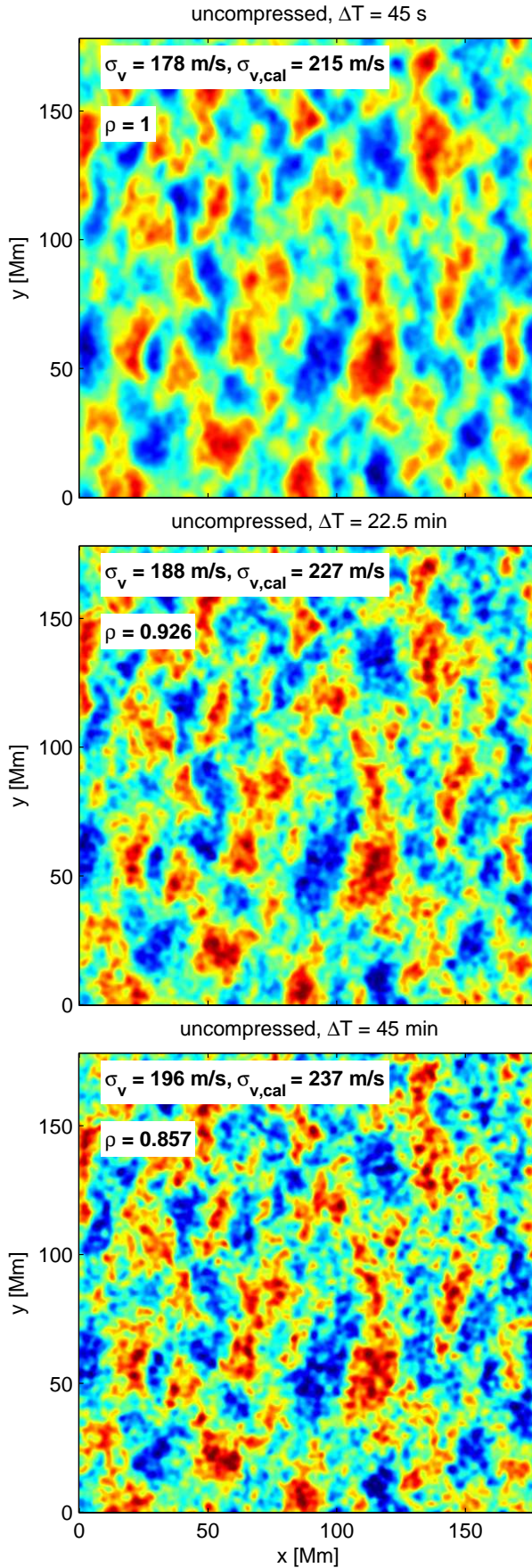


Figure 4.2: Flows in the  $x$ -direction at disk center, as determined by local correlation tracking of granulation (averaged over 24 h). Here and in Figures 4.3 and 4.4, we compare a flow map for uncompressed data with flow maps for different compression methods. In the vertical direction, we vary the time lag  $\Delta T$  between consecutive velocity maps used for averaging over 24 h (*top row*:  $\Delta T = 45$  s, *middle row*:  $\Delta T = 22.5$  min, *bottom row*:  $\Delta T = 45$  min). The flow maps shown in this figure are derived from uncompressed intensity images. Figure 4.3 is for quantized data ( $n = 2$ ) and Figure 4.4 is for JPEG compression (quality = 10). All flow maps clearly show supergranulation. We use the same scaling of the colormap for all images, corresponding to the amplitudes of the velocities before applying the calibration factor. The numbers in the images give the standard-deviation of the velocities before ( $\sigma_v$ ) and after ( $\sigma_{v,cal}$ ) applying the calibration factor to the data and the correlation with the LCT maps computed from uncompressed data ( $\rho$ ). The flow maps computed from compressed data have a higher noise level than those for uncompressed data. Quantization and JPEG lead to noise in the cross-covariances used by the LCT code and, thus, results in noise that varies on a pixel-by-pixel scale. JPEG causes a higher noise level than quantization. A large time lag between consecutive velocity maps causes granulation noise to remain visible in the image.



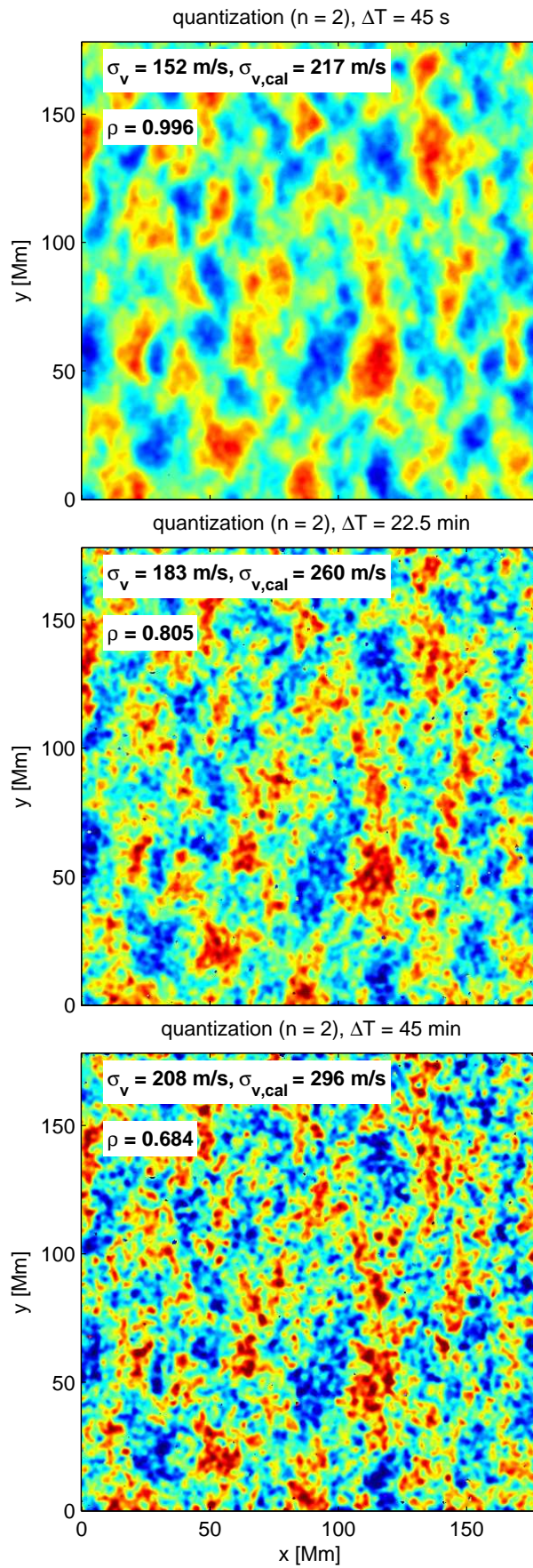


Figure 4.3: Same as Figure 4.2, but with quantization being applied to the individual images ( $n = 2$ ).

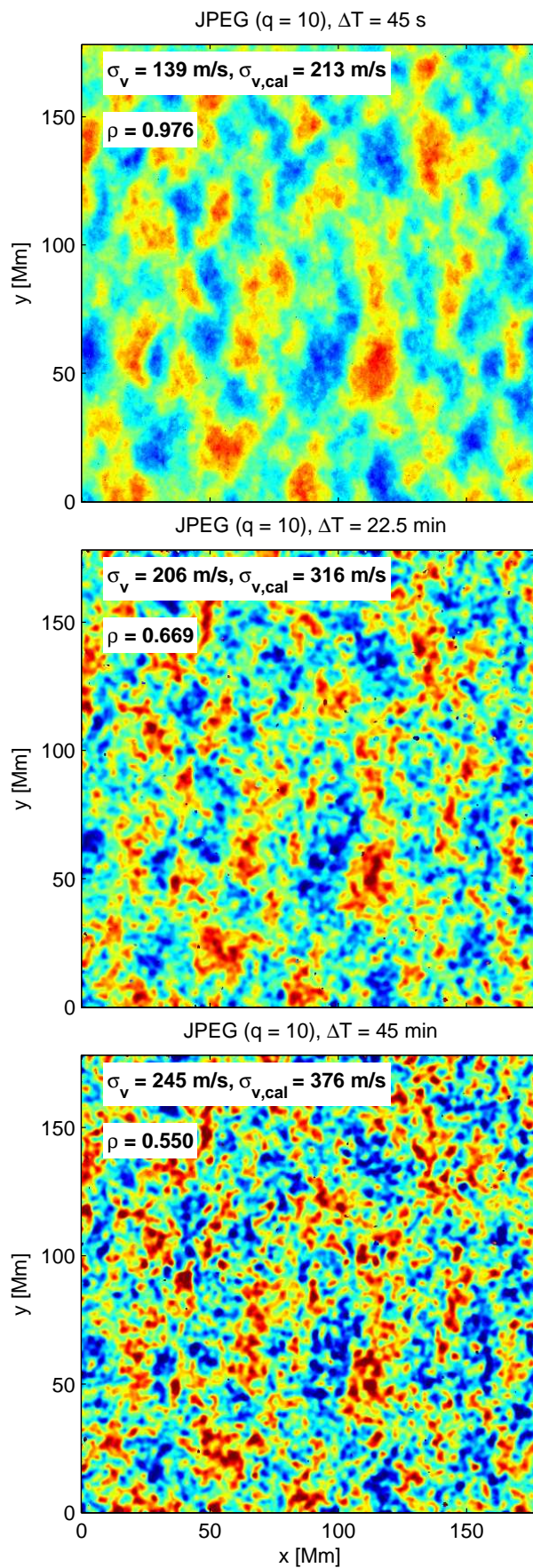


Figure 4.4: Same as Figure 4.2, but with JPEG compression being applied to the individual images (quality = 10).



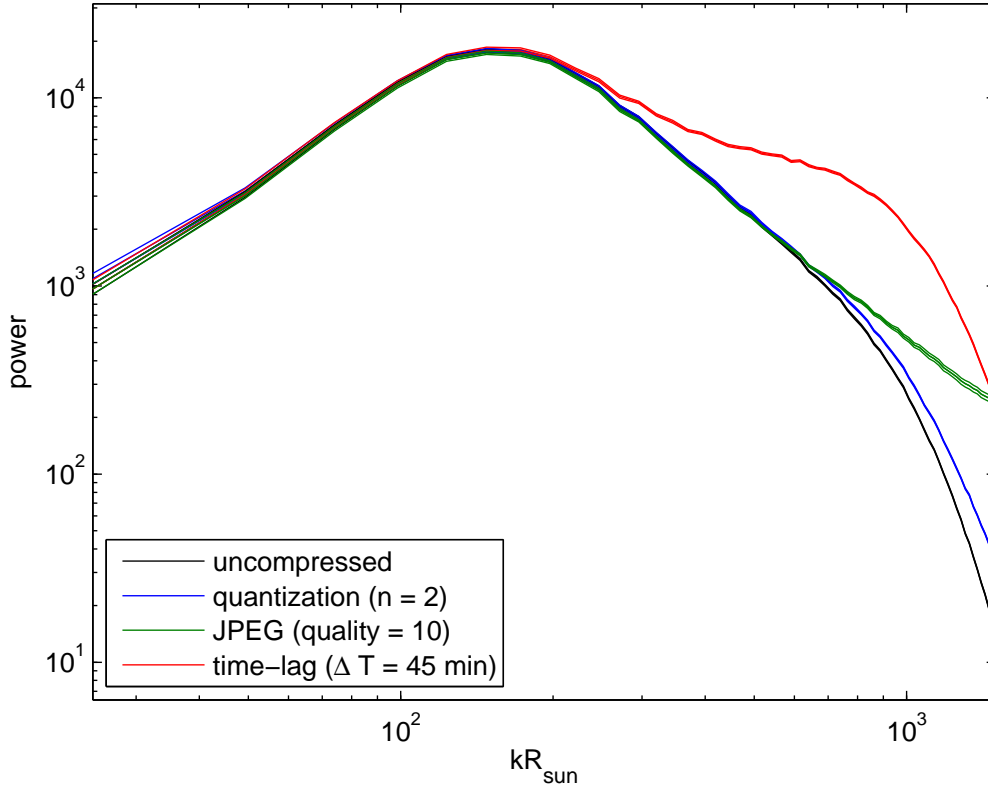


Figure 4.5: Spatial power spectra of the flow divergence at disk center (averaged power for  $T = 120$  days). *Black*: uncompressed data, *blue*: quantization ( $n = 2$ ), *green*: JPEG (quality = 10), *red*: large time lag between consecutive velocity maps ( $\Delta T = 45$  min). The thickness of the lines shows the  $1\sigma$ -scatter of the mean power. All curves exhibit a peak at  $kR_{\odot} \approx 120$  resulting from supergranulation. At high wavenumbers the additional noise introduced by the compression is visible. A large time lag between consecutive velocity maps does not remove all the noise from granulation. Quantization and JPEG lead to noise on a pixel-by-pixel scale with JPEG being worse than quantization. At supergranulation scales, the influence of quantization and JPEG compression is negligible.

deviation from perfect antisymmetry can result in a significant bias of the derived rotation rate. Figure 4.6 indeed shows a difference of  $(93 \pm 3)$  m/s between the amplitudes of the velocities in the  $x$ -direction at  $60^\circ$  longitude east and west of the central meridian. This difference is too large to be caused by rotation (difference between solar rotation and tracking rate). Unfortunately, it is not possible to determine if the shrinking-Sun effect in  $v_x$  along the central meridian exhibits an east-west asymmetry as well. This velocity is a superposition of the shrinking-Sun effect and the residual signal from differential rotation (rotation minus tracking rate). Solar rotation is constant in the east-west direction, lifting any asymmetry that might be caused by the shrinking-Sun effect.

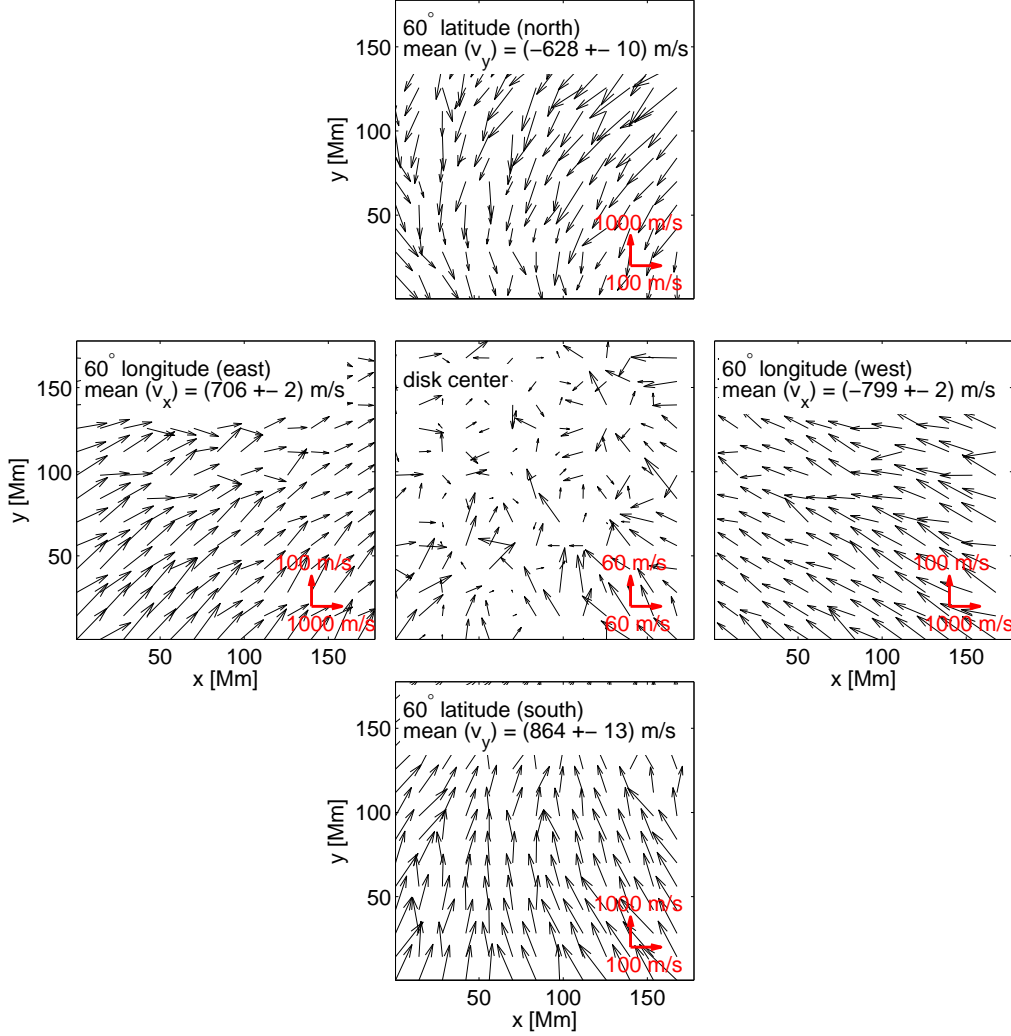


Figure 4.6: Vector velocities derived by LCT averaged over 120 days for several positions on the disk. *Top*: velocities at 60° latitude north at the central meridian, *bottom*: 60° latitude south at the central meridian, *left*: 60° longitude east at the equator, *center*: 0° latitude and 0° longitude, and *right*: 60° longitude west at the equator. The scaling of the arrows (shown in the *bottom right* in the individual images) for the plot at disk center is different to the other ones. The range of the velocities varies significantly over the disk due to different magnitudes of the shrinking-Sun effect. We amplify the flows in the azimuthal direction ( $v_x$  for 60° latitude north and south and  $v_y$  for 60° longitude east and west) by a factor of ten when displaying the arrows. The mean values of the velocities in radial direction are shown above the individual images. The amplitude of the shrinking-Sun effect is different for 60° latitude north and south. This is caused by the  $B_0$ -angle. It is mostly positive and increases from  $-4^\circ$  to  $+7^\circ$  over the course of the 120 days that we have studied. There is also a difference of  $(93 \pm 3)$  m/s in the amplitude of the shrinking-Sun effect between 60° longitude east and west. The reason for this is not understood yet. At the equator, the velocities in the  $y$ -direction caused by the shrinking-Sun effect point predominantly northwards due to the mostly positive  $B_0$ -angle.

### 4.3.3 Differential Rotation

Figure 4.7 shows differential rotation derived from LCT data. The rotation rate from Snodgrass (1984) is included for comparison. There are big differences between the rotation rate between the northern and southern hemispheres. At  $60^\circ$  latitude, the southern hemisphere rotates faster by  $(34 \pm 4)$  m/s. In the southern hemisphere, our results are in agreement with the Snodgrass (1984) rate, but in the northern hemisphere, there is an offset which increases with latitude. Of course, the Sun could have been rotating faster in the southern hemisphere than in the northern one for the studied time-period, but this difference could also be related to the shrinking-Sun effect (it is not perfectly antisymmetric along the equator, see Figure 4.6). The noise level is almost constant with latitude. Supergranulation is the main source of noise when probing differential rotation with LCT. The flow maps all have the same size ( $178 \times 178$  Mm) and so, they contain the same number of supergranules and give a similar noise level.

Figures 4.7 and 4.8 also demonstrate how the various compression methods affect measurements of differential rotation. In order to save computation time we only computed the rotation rate for positive latitudes from the compressed data. The time lag method exhibits a large bias that increases linearly with latitude. Quantization and JPEG compression are in good agreement with the uncompressed case. We also include data centered at  $10^\circ$  latitude, overlapping with the intensity images centered at  $0^\circ$  and  $20^\circ$ . The overlapping parts are in good agreement, suggesting that measuring rotation and compression are not affected by projecting the intensity images on a cartesian coordinate system.

### 4.3.4 HMI *P*-Angle Oscillations

The large bias of the time lag method is caused by oscillations of the *P*-angle of the HMI instrument that, as far as we know, have not been reported before. The roll angle of the SDO spacecraft relative to the rotation axis of the Sun is given by the keyword *CROTA2* and is saved for each individual filtergram obtained by HMI (3.75 s cadence). This keyword shows oscillations with a fundamental period of 27.8 mHz and several harmonics. These oscillations are also present in the *CROTA2* keyword of the 45 s cadence computed continuum intensity images. Due to aliasing the dominant frequency appears at 5.56 mHz and two harmonics are at 2.78 and 8.33 mHz ( $\sim 10^{-4}$  deg rms in total). The origin of these oscillations is currently unclear. Such a periodic shift of the intensity images affects, of course, the derived LCT velocities. For data along the central meridian the velocities in the *x*-direction and for the cubes along the equator in the velocities in the *y*-direction exhibit oscillations at the same frequencies as the *CROTA2* keyword. The amplitude of the oscillations in the LCT data can be estimated from a simple model of a disk rotating around its center with an angular velocity given by the changes of the *CROTA2* keyword. However, the amplitudes in the power spectrum predicted by this simple model do not match exactly with the observations in LCT (about 10% difference). In theory, these oscillations should be removed by the tracking code. This is not the case here, indicating that the *CROTA2* keyword does not describe the *P*-angle oscillations correctly.

The oscillations have a short period and mostly cancel out when averaging a sufficient amount of data. However, if the time lag  $\Delta T$  between consecutive velocity images is in resonance with these oscillations (as it is the case for  $\Delta T = 30$  min, corresponding to a

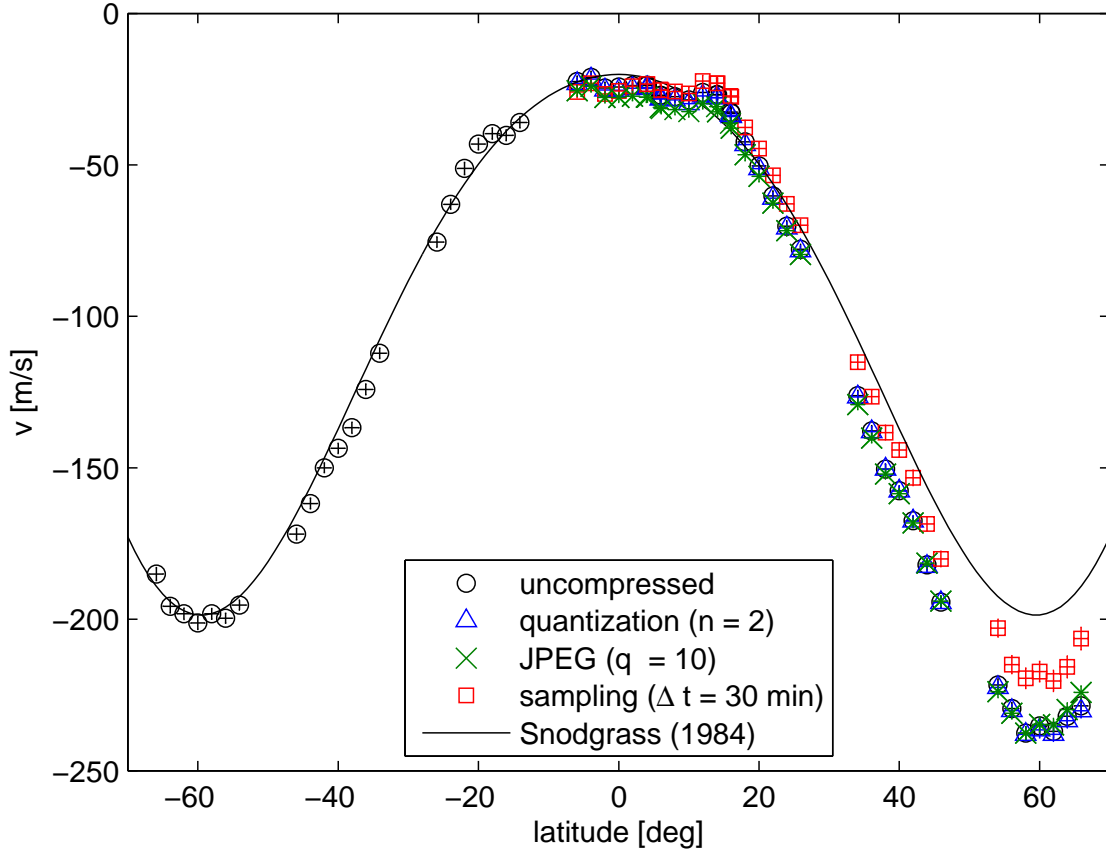


Figure 4.7: Differential rotation relative to solid body rotation determined using LCT on uncompressed and compressed data ( $T = 120$  days). *Black circles*: uncompressed data, *blue triangles*: quantization ( $n = 2$ ), *green crosses*: JPEG (quality = 10), *red squares*: large time lag between consecutive velocity maps ( $\Delta T = 30$  min). The solid curve is the Snodgrass (1984) rotation rate which is used by the tracking code. In order to save computation time, we only computed the rotation rate for positive latitudes with compressed data. In the northern hemisphere, there is a large (up to  $\sim 34$  m/s) offset between the Snodgrass (1984) rate and the velocities derived using LCT. Its origin is unknown. The error bars are smaller than the symbol size.

frequency of 0.556 mHz), this effect can lead to an offset of the order of up to 20 m/s in the flow maps averaged over 24 h (as observed in Figure 4.7). In Figure 4.9, we tried to remove these oscillations by using the simple model described above. The remaining bias is comparable to the noise level. There might still be some residual signal, though, because the amplitude of the effect predicted from the *CROTA2* keyword does not match exactly with that from the LCT velocities (as discussed above). In the remainder of this paper, we only show results after applying this correction.

#### 4.3.5 Influence of Compression on Differential Rotation

Compression affects LCT in several ways. It could potentially change the noise level, systematic errors, and also the signal from solar rotation itself. Determining the impact

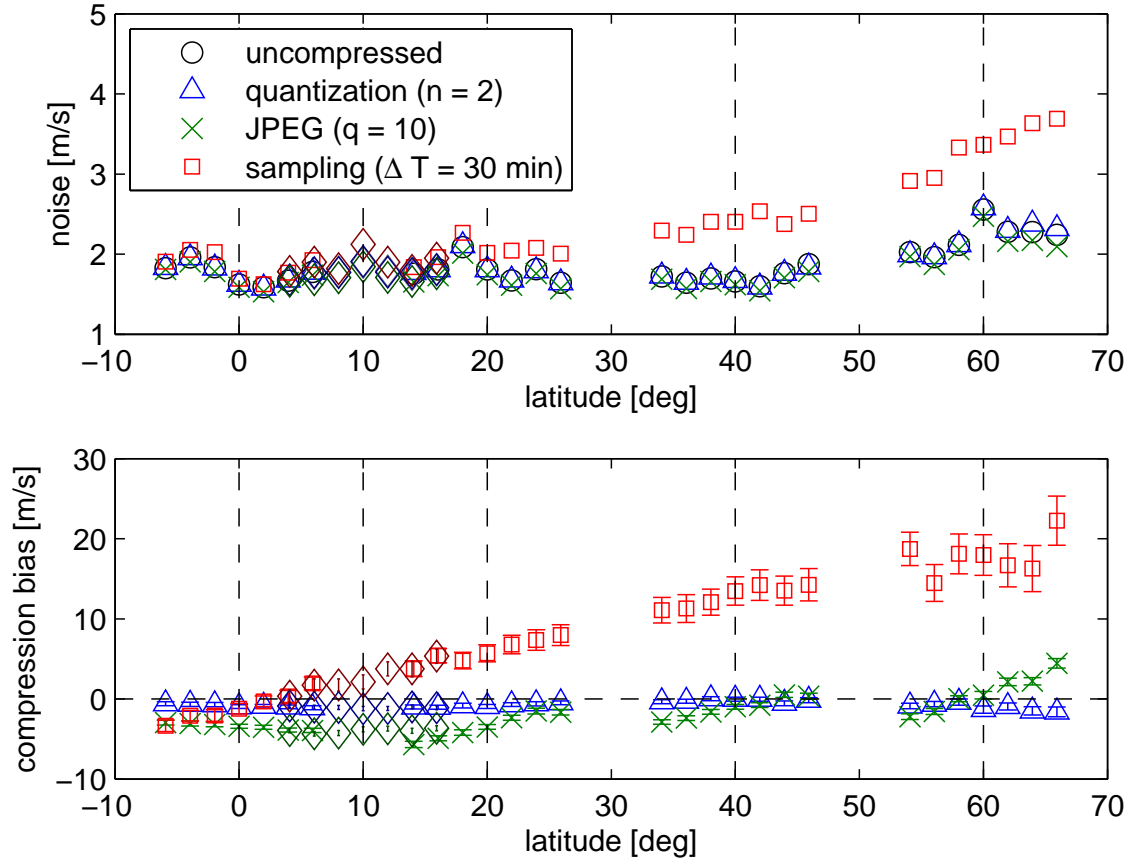


Figure 4.8: *Top*: noise level and *bottom*: compression bias (difference compressed minus uncompressed data) as a function of latitude for the differential rotation curve shown in Figure 4.7. The vertical dashed lines show the centers of the tracked and remapped intensity images. The *diamonds* show the noise and bias derived from intensity data centered at  $10^\circ$  latitude, overlapping with the data centered at  $0^\circ$  and  $20^\circ$  latitude. The time lag method suffers from a large compression bias that increases linearly with latitude. Quantization and JPEG are in good agreement with the uncompressed case.

on the noise level is straightforward, but it is hard to measure the impact on the expectation value of the rotation velocities. It is only possible to determine the compression bias, defined as the difference compressed minus uncompressed. This compression bias is also affected by changes of the noise or systematic errors. For all compression methods, the expectation value of the bias does not depend on the observing time. Noise and compression bias for differential rotation are shown in Figure 4.9.

The time lag method suffers mostly from the oscillations of the  $P$ -angle of HMI (as discussed in the previous section). In addition, the noise level is increased significantly due to the granulation noise. This probably also causes the remaining bias of the time lag method (after correcting for the  $P$ -angle oscillations). Quantization and JPEG compression have almost no influence on the noise level. The noise in the individual flow maps is on very small spatial scales and is removed almost entirely when averaging over longitude. The rotation velocities computed from these data are also strongly correlated with the uncompressed data, indicating that the bias is not caused by changes of the noise.

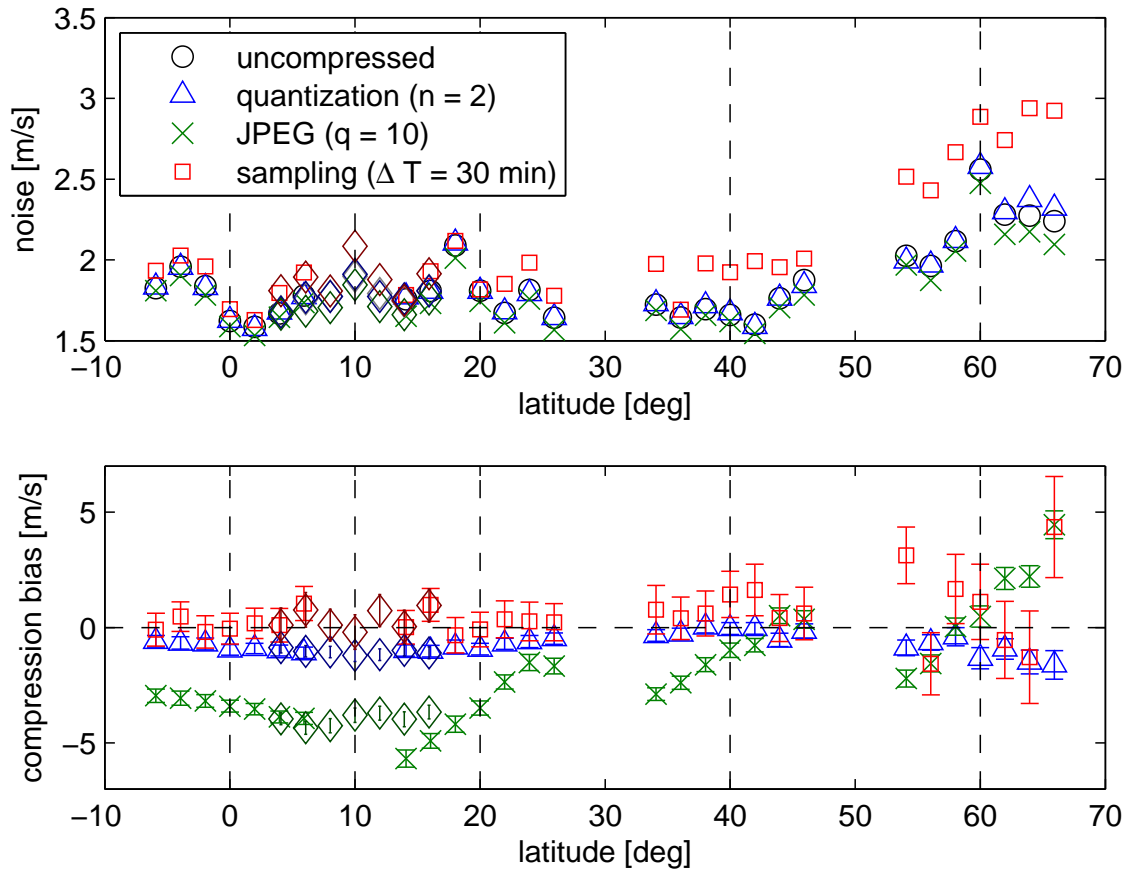


Figure 4.9: Same as Figure 4.8, but with a simple correction method being applied to the time lag method that is based on the *CROTA2* keyword of the HMI data. More details are described in the text. This correction removes most of the bias and slightly reduces the noise level. The remaining bias of the time lag method is comparable to the noise level. The time lag method suffers from a higher noise level than the uncompressed data, since granulation noise is still present in the data. Quantization and JPEG compression have almost no influence on the noise level. The compression bias for quantization is comparable with the noise level. JPEG compression exhibits a strong bias that varies with latitude. Note that the ranges of the y-axes are different to those in Figure 4.8.

In most cases, the compression bias introduced by quantization and JPEG compression is negative. Quantization and JPEG compression enhance the shrinking-Sun effect. If the deviation of the uncompressed case from the Snodgrass (1984) rate is really caused by the shrinking-Sun effect, this offset will be amplified by compression, leading to a negative compression bias. Some of the bias of the JPEG method seems to be originating from inaccuracies of the calibration. The compression bias at around  $15^\circ$  latitude is different for the data centered at  $10^\circ$  and  $20^\circ$  latitude. In addition, the bias for JPEG compression exhibits a linear trend within the individual datacubes.

The amount of additional noise and the compression bias depend on the method and the parameters of the compression. Figures 4.10 and 4.11 show compression bias and noise level averaged over latitude for all values of  $n$  and  $q$  and several time lags  $\Delta T$ . Quantization causes a much smaller compression bias than JPEG. Only very low values of  $n$  lead to a significant bias. Both quantization and JPEG compression have almost no impact on the noise level. These methods introduce noise on small spatial scales, which is not relevant when measuring rotation. The time lag method increases the noise level and bias significantly.

## 4.4 Discussion and Conclusion for Chapter 4

The results of this paper suggest that LCT is robust to data compression when using quantization. However, we only looked at differential rotation and supergranulation, which does not need to be representative for LCT in general. Measurements that rely more on small spatial scales might be affected by an increased noise level to a higher degree. Also, we tested the influence of compression using tracked and remapped intensity images, applying it on raw full disk images might give different results.

Figures 4.10 and 4.11 show that the file size can be reduced by a factor of ten relative to the size of the HMI raw images without significant bias ( $< 0.5$  m/s) or increased noise (no increase within the error bars). The influence of the compression bias relative to the noise level depends on the observing time. The expectation value of the bias is constant with time but the noise level decreases with increasing observing time. Both for individual flow maps and for differential rotation, quantization is the best compression method. JPEG compression leads to a higher noise level and it decreases the sensitivity of LCT more than quantization. JPEG compression removes information about small spatial scales, including the granulation signal used by LCT (see Löptien et al. 2014a). Quantized images on the other hand clearly show granulation, even for small  $n$ . Löptien et al. (2014a) found that JPEG was the best compression method tested for time-distance helioseismology. This suggests applying different compression methods to the Dopplergrams used by local helioseismology and the continuum images used by LCT. The time lag method is very sensitive to changes of the geometry of the data. It can only be used if the pointing and roll angle of the instrument are both extremely stable with time. When deciding to use this method, it depends on the constraints of the observations which time lag  $\Delta T$  to select. If both telemetry and observing time are fixed, the time lag should be as short as the telemetry allows, because this decreases the granulation noise more efficiently. If there is only a limit on telemetry and the image geometry is well known, observations with a large time lag for a long observing time are preferable. This allows

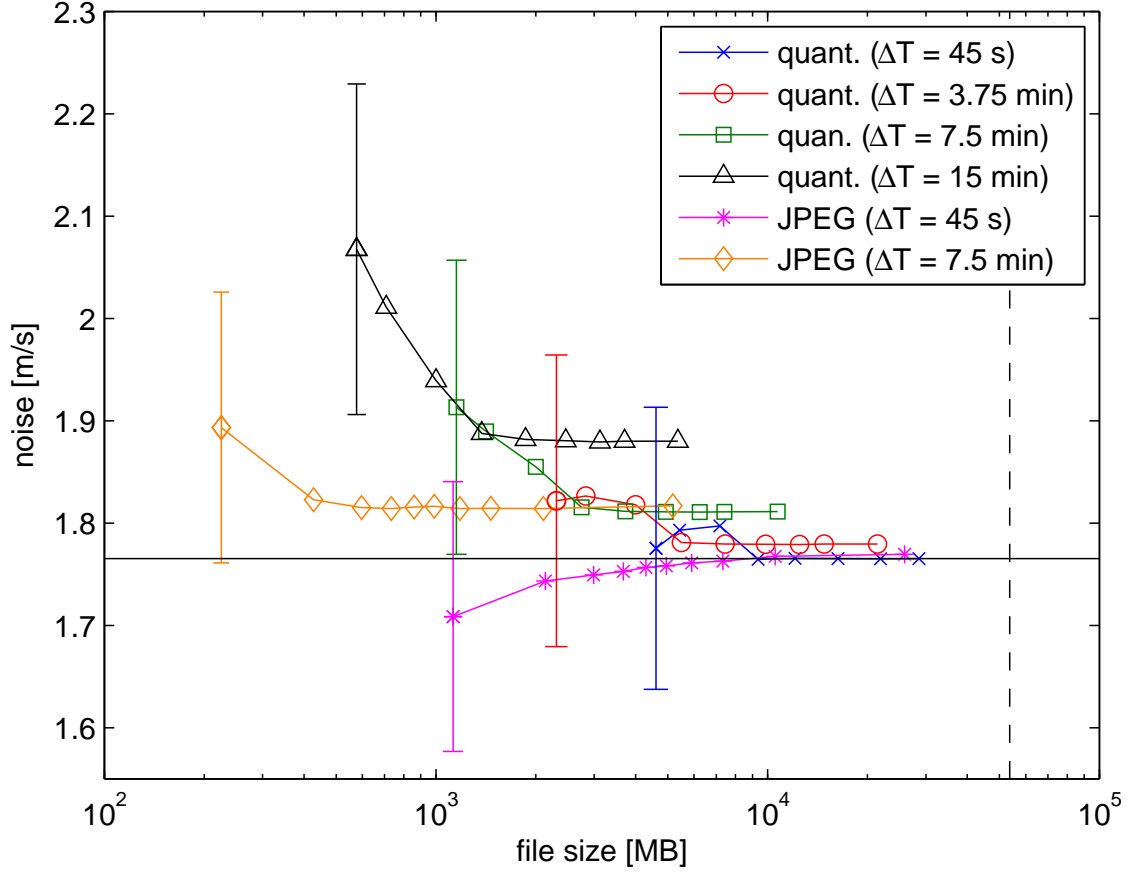


Figure 4.10: Noise level averaged over latitude (not including the data centered at  $10^\circ$  latitude) as a function of file size for different compression methods. The different curves correspond to the different compression methods (quantization and JPEG) and to data with different time lags  $\Delta t$  between consecutive pairs of images. *Blue crosses*: quantization with  $\Delta T = 45$  s, *red circle*: quantization with  $\Delta T = 3.75$  min, *green squares*: quantization with  $\Delta T = 7.5$  min, *black triangles*: quantization with  $\Delta T = 15$  min, *purple asterisks*: JPEG compression with  $\Delta T = 45$  s, and *orange diamonds*: JPEG compression with  $\Delta T = 7.5$  min. Along the curves, we vary the number  $n$  of possible values for the intensity for quantization or the quality factor  $q$  for JPEG compression. The file size represents the file size necessary for measuring differential rotation as in Figure 4.7 ( $T = 120$  days). The horizontal line represents the noise level averaged over latitude for uncompressed data ( $T = 120$  days). The vertical line shows the size of the uncompressed data (assuming a size of 7.1 bits per pixel for the HMI raw data). For clarity, we only show one error bar for each curve. We only show results for short time lags. Larger time lags cause significantly higher compression bias and noise level.



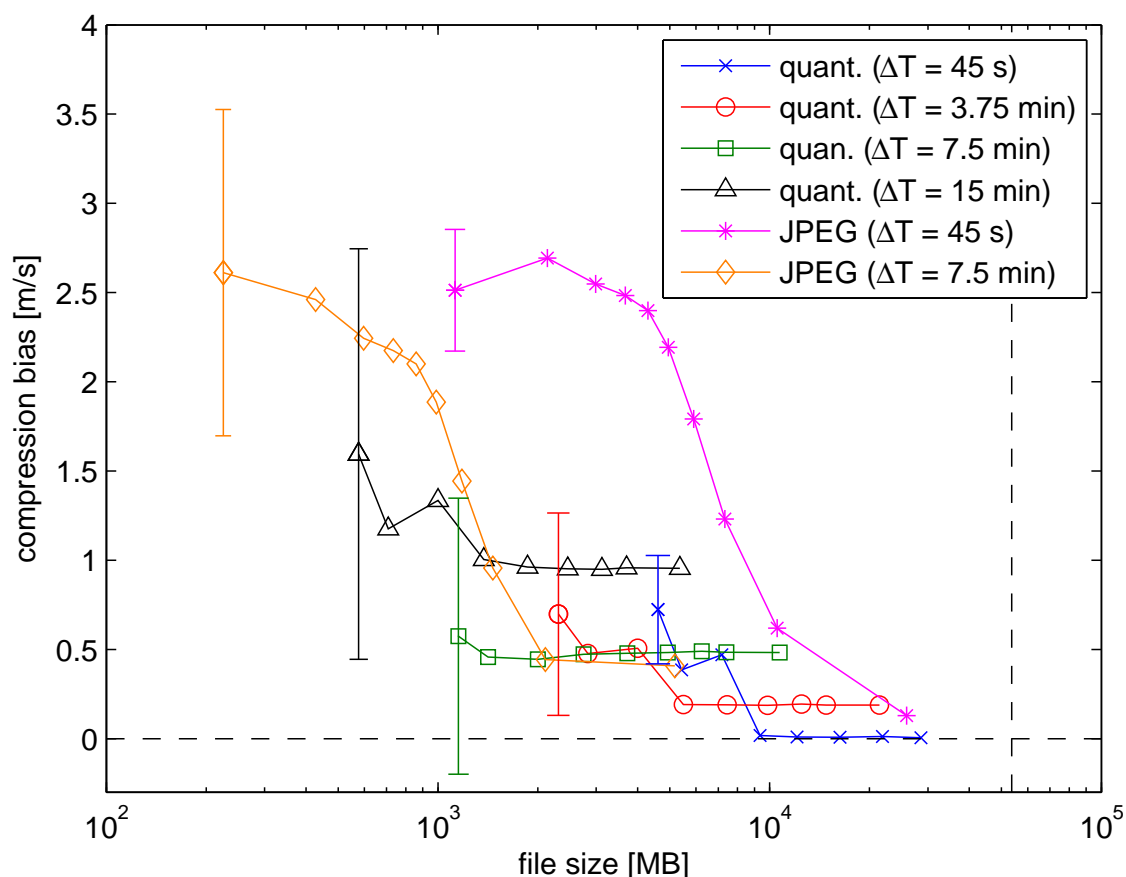


Figure 4.11: Same as Figure 4.10, but showing the compression bias averaged over latitude (not including the data centered at  $10^\circ$  latitude) as a function of file size for different compression methods.

sampling more supergranules and thus, reduces the supergranulation noise further.

It might be possible to remove the compression bias from the flow velocities if its origin were understood. At least part of it could be caused by the shrinking-Sun effect. This effect and its coupling with compression might be responsible for the large differences between the northern and southern hemisphere in our rotation velocities. Our results motivate a detailed study of the origin of the shrinking-Sun effect and its observed asymmetry along the equator (see Figure 4.6). One possibility would be generating synthetic data from simulations of solar surface convection. Understanding the shrinking-Sun effect is also a requirement for applying LCT to *Solar Orbiter* data. Probing flows near the poles, even with *Solar Orbiter*, means observing far from disk center, where systematic effects become more important. In addition, the sensitivity of other feature tracking algorithms like the coherent structure tracking (CST) code (Rieutord et al. 2007, Tkaczuk et al. 2007, Roudier et al. 2013) regarding systematic errors and compression needs to be tested.

In principle, it is also possible to run an LCT algorithm onboard the spacecraft and to transfer only the derived flow maps. This would decrease the required telemetry significantly, but would require an extremely good calibration of the instrument. The geometry of the data would have to be well known. Any systematic errors present in the continuum intensity maps would also affect the LCT flow maps.

A detailed study of the influence of data compression on LCT is not only important for *Solar Orbiter*, but also for other planned missions, such as L5 or SAFARI (see Sekii et al. 2015, for a review on concepts for future missions).

## 5 Discussion and Outlook

### 5.1 Feasibility of Helioseismic Science Goals

There are four main challenges for performing helioseismology with *Solar Orbiter*: low telemetry, short observing times, systematic errors of the analysis methods, and systematic errors caused by the PHI instrument. In this thesis, I addressed two of them. I evaluated the impact of data compression and looked at the influence of the PHI instrument on the observation of solar oscillations.

The most important result of this thesis is the surprisingly good performance of lossy data compression. Although this needs to be tested in more detail, the results obtained in this thesis suggest that telemetry will not be the limiting factor for helioseismology with *Solar Orbiter*. The compression algorithm of choice depends on the analysis method. For time-distance helioseismology, JPEG is the best method tested, for LCT, quantization provides the better results.

More worrying than the telemetry are the probably short observing times for helioseismology. Although it is being discussed to have an extended observing run for helioseismology, it will be significantly shorter than what is possible with current missions. The orbit of *Solar Orbiter* does not allow extended observations from a fixed vantage point, e.g., the spacecraft will not be longer at solar latitudes greater than  $30^\circ$  for more than 21 consecutive days. The resulting noise level will probably be too high for achieving all of the science goals to be addressed by the mission.

Another big concern are systematic errors. Even with the inclined orbit of *Solar Orbiter*, probing the polar regions of the Sun requires observations close to the limb. Here, systematic errors like the shrinking-Sun effect or the phase-shift in time-distance helioseismology become more and more important, especially when measuring the meridional flow. As shown in Chapter 4, the shrinking-Sun effect dominates the LCT signal when measuring away from disk center.

Systematic errors originating from the PHI instrument are also going to be a problem for *Solar Orbiter*. According to the tests performed in Section 2, PHI will be suitable for measuring solar oscillations. However, I only tested the influence of different instrumental effects for an observation from a fixed vantage point. Some properties of PHI, such as the point spread function, might change with time, due to the orbit. This could lead to time-dependent systematic errors in the data. In any case, PHI will need to be calibrated extremely well.

To summarize, the *Solar Orbiter* mission is not designed perfectly for helioseismology in several aspects. The mission will cover a broad range of topics in solar physics, with helioseismology being only one of them. This requires compromises in the design

of the mission. However, even if *Solar Orbiter* will not fulfill all of the science goals involving helioseismology, it will still be a very important mission for helioseismology. No other space mission has been approved yet that could address the science questions to be answered by *Solar Orbiter*. There are concepts for such missions, e.g., L5 or SAFARI, and if they get selected at some point, they will greatly benefit from the experience gained from *Solar Orbiter*.

## 5.2 Outlook: Probing Differential Rotation with Time-Distance Helioseismology

In an ongoing project, I evaluate the influence of data compression when measuring differential rotation using time-distance helioseismology with the f-mode. As can be seen in Figure 5.1, LCT always provides a slower rotation rate than my preliminary results for time-distance helioseismology. The difference in rotation between the northern and southern hemispheres and the torsional oscillations at  $15^\circ - 20^\circ$  are also visible in the time-distance data. At all latitudes, LCT has a lower noise level than time-distance helioseismology.

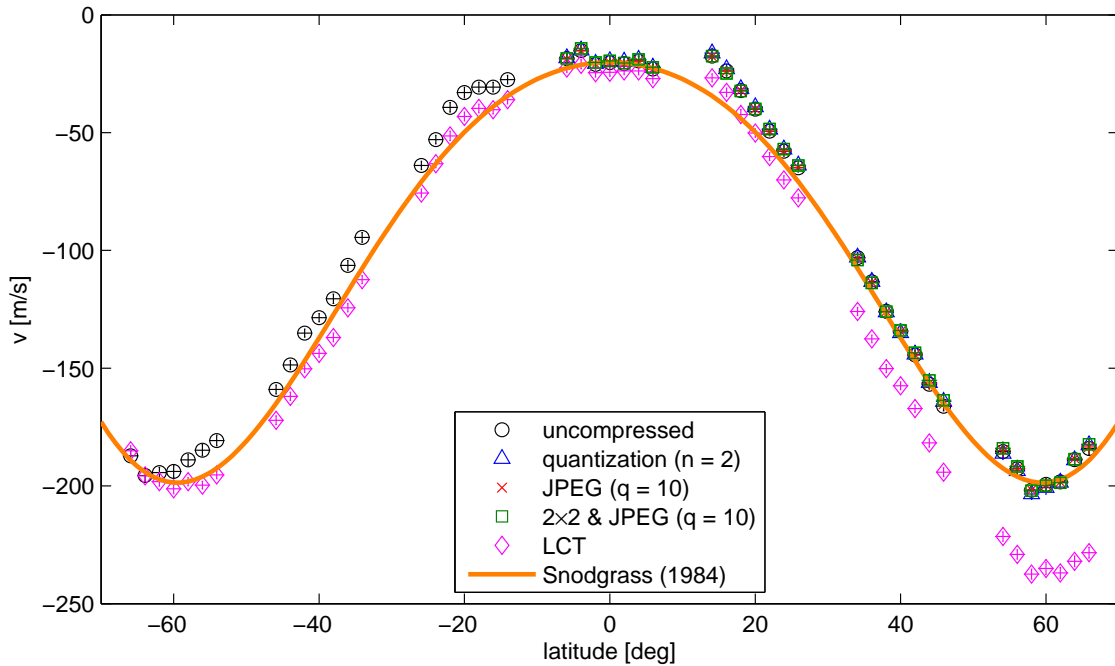


Figure 5.1: Differential rotation relative to solid body rotation determined using time-distance helioseismology (f-mode,  $T = 120$  days) for uncompressed and compressed data. *Black circles*: uncompressed data, *blue triangles*: quantization ( $n = 2$ ), *red crosses*: JPEG (quality = 10), *green squares*:  $2 \times 2$  subsampling and JPEG compression (quality = 10). For comparison, the results from LCT (*purple diamonds*) and the Snodgrass (1984) rotation rate (*orange curve*) are shown as well. The error bars are smaller than the symbol size.

Figure 5.1 also shows rotation derived from Dopplergrams that have been compressed

using quantization ( $n = 2$ ), JPEG compression ( $q = 10$ ), or  $2 \times 2$  subsampling followed by JPEG compression ( $q = 10$ ). These are almost indistinguishable from the uncompressed case.

The influence of compression is evaluated in more detail in Figure 5.2. All tested methods do not increase the noise level significantly, and in most cases, the compression bias is negligible, as well (smaller than the noise level). These preliminary results suggest that probing differential rotation with time-distance helioseismology is robust against data compression.

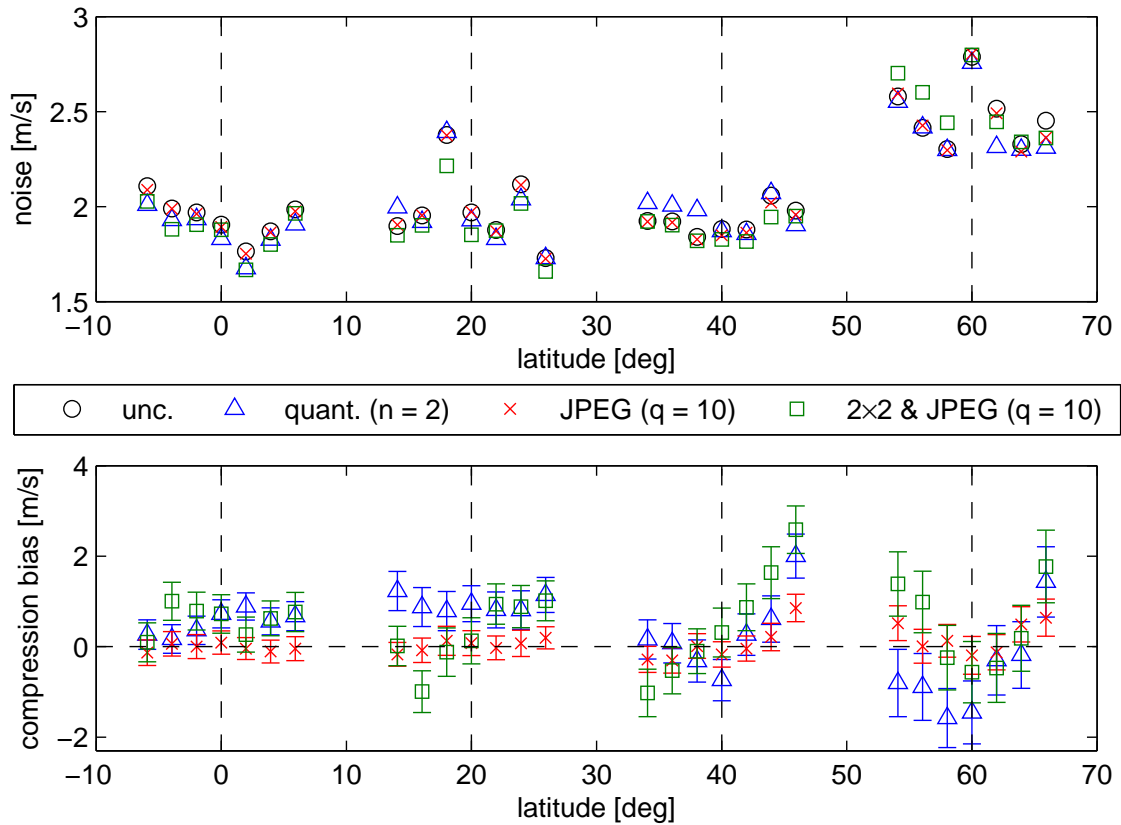


Figure 5.2: Influence of compression on measuring rotation with time-distance helioseismology. *Top*: Noise in differential rotation as a function of latitude ( $T = 120$  days), *bottom*: Compression bias, defined as the difference compressed minus uncompressed, as a function of latitude. *Blue triangles*: quantization ( $n = 2$ ), *red crosses*: JPEG (quality = 10), *green squares*:  $2 \times 2$  subsampling and JPEG compression (quality = 10). The vertical dashed lines show the centers of the tracked and remapped cubes.

### 5.3 Outlook: Modeling the Shrinking-Sun Effect

As discussed in Chapter 4, the shrinking-Sun effect in LCT is going to be a problem for *Solar Orbiter*. It severely affects measurements at high latitudes, especially of the meridional flow. In this Section, I show that one promising avenue for understanding the shrinking-Sun effect and other systematic errors is using synthetic data derived from

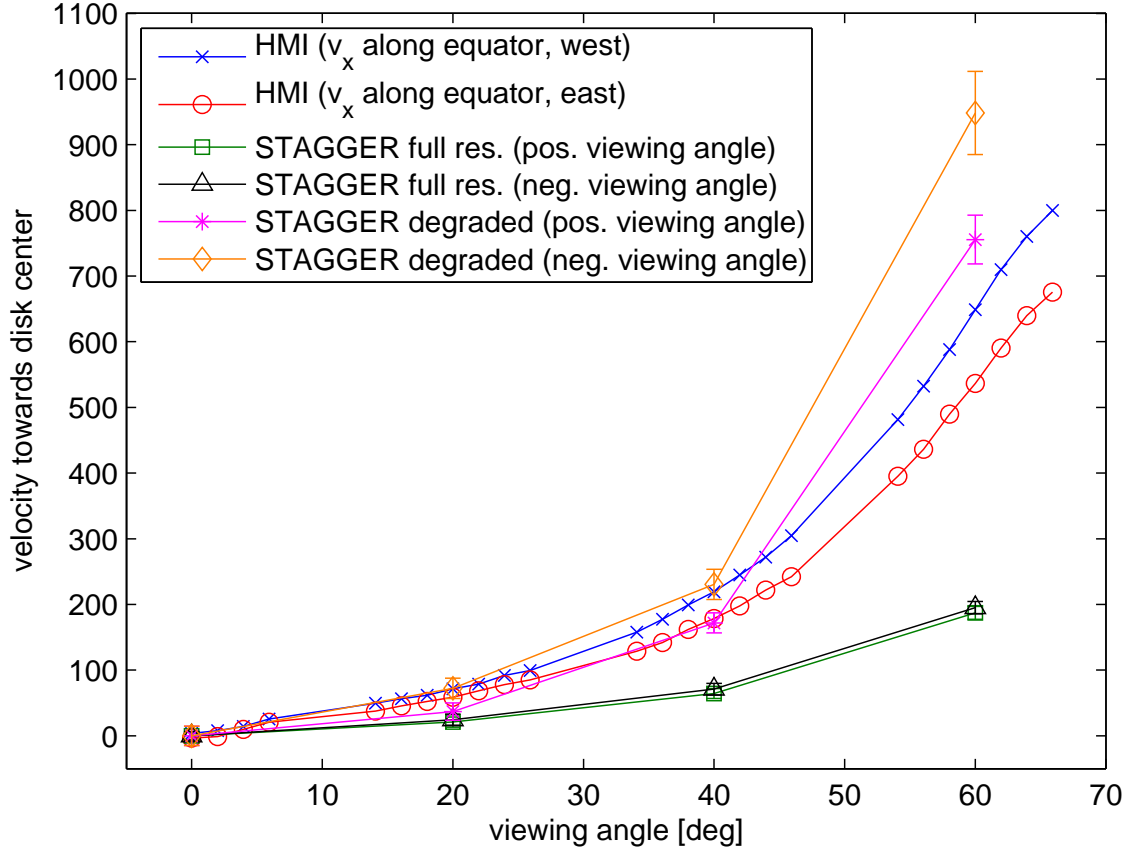


Figure 5.3: Velocity towards disk center determined using LCT as a function of distance to disk center averaged over  $x$  and  $y$ . I compare velocities derived from HMI continuum intensity images with two sets of synthetic data derived from the STAGGER simulations. The HMI data consists of the velocity in the  $x$ -direction derived using LCT along the equator for longitudes west (*blue crosses*) and east (*red circles*) of the central meridian. The first set of synthetic data consists of STAGGER data full full spatial resolution for positive (*green squares*) and negative (*black triangles*) viewing angles. The second set of synthetic data consists of STAGGER data convolved with a point spread function (psf) corresponding to the psf of HMI and resampled to the pixel size of the tracked and remapped HMI data (348 km) for positive (*purple stars*) and negative (*orange diamonds*) viewing angles. The error bars of the HMI data are smaller than the symbol size. See the text for more details about the synthetic data.

simulations of solar surface convection. The advantage of this approach is that synthetic data corresponding to different observing conditions can easily be compared because the properties of the underlying simulations are the same in each case.

Figure 5.3 compares the magnitude of the shrinking-Sun effect as a function of distance to disk center for HMI continuum intensity and synthetic data. The HMI continuum images are the same dataset as used in Chapter 4, but include also intensity images centered at  $20^\circ$  and  $40^\circ$  east and west of the central meridian along the equator. The synthetic data consist of a time-series of continuum intensity images generated from the STAGGER simulations (size:  $48 \times 48$  Mm) using the SPINOR code with different values of the

viewing angle  $\theta$  ( $-60^\circ, -40^\circ, -20^\circ, 0^\circ, +20^\circ, +40^\circ, +60^\circ$ ). Results for two different sets of synthetic data are shown. In the first case, I run the FLCT code on synthetic continuum intensity map with full spatial resolution. In the second case, the continuum images are convolved with a point spread function (psf) corresponding to the psf of HMI (approximated as an Airy-function) and resampled to the pixel size of the tracked and remapped HMI data (348 km) before running the FLCT code on them. For each value of  $\theta$ , I generate 12 pairs of consecutive continuum intensity images with the pairs being separated by 30 min in order to save computation time. Both synthetic datasets exhibit the shrinking-Sun effect. If the psf is accounted for, the mean velocities of the synthetic data generated from the STAGGER simulations match well with those of the observed data.

The shrinking-Sun effect is already present in the synthetic data without psf. For a fixed value of  $\theta$ , it causes an almost constant offset relative to the data for  $\theta = 0^\circ$  in the velocity maps pointing towards disk center. Apart from this, there is no difference in the flow maps for different values of  $\theta$ . The magnitude of this effect increases monotonically with  $\theta$ . Hence, the origin of the shrinking-Sun effect has to be related directly to the change of the viewing angle. In addition, the velocity caused by the shrinking-Sun effect is always directed towards disk center, it does not only depend on the distance to disk center, but also on the direction. Hence, the shrinking-Sun effect must be connected to some asymmetry between the directions center-limb and limb-center imposed by the viewing angle. The only asymmetry like this in the synthetic data is that the path of the light in the photosphere is inclined when  $\theta > 0$ . This causes the side of granules pointing towards disk center to be brighter than the one directed towards the limb (hot-wall effect).

The shrinking-Sun effect in the HMI continuum images is not symmetric around the equator (see Figure 5.3). As already discussed in Chapter 4, it is much stronger for data west of the central meridian than for data east of the central meridian ( $112 \pm 4$  m/s difference at  $60^\circ$  longitude). Such an asymmetry is also observed in the synthetic data between positive and negative values of  $\theta$ . This asymmetry is weak for the data with full spatial resolution ( $8 \pm 14$  m/s difference between  $\theta = +60^\circ$  and  $\theta = -60^\circ$ ), but increases significantly when applying the psf ( $190 \pm 80$  m/s difference between  $\theta = +60^\circ$  and  $\theta = -60^\circ$ ). The origin of this asymmetry is currently unclear. Further work is required, using both the STAGGER simulations and additional observations, e.g., data provided by MDI.

## 5.4 Future Work

The results of this thesis are just a first step in preparing for the *Solar Orbiter* mission. A lot of work still needs to be done.

Each science goal of the mission requires the development of a detailed observing strategy. This involves accurate noise estimates, comparing the efficiency of different compression algorithms and estimating the influence of systematic errors. Here, I focused on time-distance helioseismology with the f-mode and LCT. More methods for measuring flows have to be tested, such as supergranulation tracking or coherent structure tracking (CST).

The most important issue to be addressed is compression. The results presented in this thesis are not yet sufficient for deciding on a strategy for compressing PHI data, because the good performance of compression in the tests performed in this thesis does not need

to be the case for helioseismology in general. The impact of the various compression methods depends on the selected analysis method (e.g., time-distance helioseismology or LCT) and also on the science question to be addressed. Measurements with a low S/N are probably much more sensitive to compression. In addition, I tested the influence of compression on HMI data that have already been tracked and remapped. The raw data obtained by *Solar Orbiter* is probably more affected by compression, especially close to the limb where foreshortening becomes significant. Extensive compression can only be applied if the origin of the compression bias, as observed e.g., when probing differential rotation, is understood. In addition, the final decision for a particular compression method needs to account for the capabilities of the hardware and software onboard *Solar Orbiter*.

Understanding systematic errors will be one of the most important steps in preparing for helioseismology with *Solar Orbiter*. As shown in the previous section, a promising approach for this is detailed modeling using simulations of solar surface convection. Combined observations of *Solar Orbiter* and another instrument of the same region on the Sun will also help in the understanding of systematic errors.

*Solar Orbiter* also demands the development of new methods for analyzing helioseismic data. PHI will be the first helioseismic instrument that observes the Sun from a vantage point that changes with time. The resulting changes of viewing angle and spatial resolution have to be accounted for. Especially stereoscopic helioseismology requires the development of new methods, i.e. ways to probe the deep solar interior using time-distance helioseismology or holography. In order to derive observing strategies for stereoscopic helioseismology, one has to obtain an estimate of the noise level, e.g., by using global simulations of solar oscillations.

Helioseismic observations with *Solar Orbiter* will require a significant amount of telemetry and observing time. This needs to be negotiated with the teams of the other instruments onboard the spacecraft. Certainly, it will be of great help if there is a detailed strategy for addressing the various helioseismic science goals and if it is clear what can be learned about the Sun from the data to be delivered by the mission. Addressing all of the issues listed here involves a large amount of work. It needs to be decided on how much effort should be spent on preparing for the *Solar Orbiter* mission given the expected capabilities of the mission in helioseismology. Of course, such a study would also be of great help for other planned helioseismology mission, such as SAFARI or L5.



# Bibliography

- AKI, K. (1965) A Note on the Use of Microseisms in Determining the Shallow Structures of the Earth's Crust. *Geophysics*, **30**, 665.
- APPOURCHAUX, T. (2011) PHI Filtergraph Requirements and Specifications. *Technical report*.
- BALDNER, C. S. & SCHOU, J. (2012) Effects of Asymmetric Flows in Solar Convection on Oscillation Modes. *ApJ*, **760**, L1.
- BARNES, G., BIRCH, A. C., LEKA, K. D. & BRAUN, D. C. (2014) Helioseismology of Pre-emerging Active Regions. III. Statistical Analysis. *ApJ*, **786**, 19.
- BARTHOL, P., GANDORFER, A., SOLANKI, S. K., SCHÜSSLER, M., CHARES, B., CURDT, W., DEUTSCH, W., FELLER, A. ET AL. (2011) The Sunrise Mission. *Sol. Phys.*, **268**, 1–34.
- BECKERS, J. M. (1998) On the Effect of Narrow-Band Filters on the Diffraction Limited Resolution of Astronomical Telescopes. *A&AS*, **129**, 191–194.
- BELLOT RUBIO, L. R. & BORRERO, J. M. (2002) Iron Abundance in the Solar Photosphere. Application of a Two-Component Model Atmosphere. *A&A*, **391**, 331–337.
- BIRCH, A. C., BRAUN, D. C., LEKA, K. D., BARNES, G. & JAVORNIK, B. (2013) Helioseismology of Pre-emerging Active Regions. II. Average Emergence Properties. *ApJ*, **762**, 131.
- BRAUN, D. C. & BIRCH, A. C. (2008) Prospects for the Detection of the Deep Solar Meridional Circulation. *ApJ*, **689**, L161–L165.
- BRAUN, D. C., BIRCH, A. C., BENSON, D., STEIN, R. F. & NORDLUND, Å. (2007) Helioseismic Holography of Simulated Solar Convection and Prospects for the Detection of Small-Scale Subsurface Flows. *ApJ*, **669**, 1395–1405.
- CHRISTENSEN-DALSGAARD, J. (2003) Lecture Notes on Stellar Oscillations.
- CHRISTENSEN-DALSGAARD, J., DUVAL, JR., T. L., GOUGH, D. O., HARVEY, J. W. & RHODES, JR., E. J. (1985) Speed of Sound in the Solar Interior. *Nature*, **315**, 378–382.
- CLAVERIE, A., ISAAK, G. R., MCLEOD, C. P., VAN DER RAAY, H. B. & CORTES, T. R. (1979) Solar Structure from Global Studies of the 5-minute Oscillation. *Nature*, **282**, 591–594.

- COUVIDAT, S. & BIRCH, A. C. (2009) Helioseismic Travel-Time Definitions and Sensitivity to Horizontal Flows Obtained from Simulations of Solar Convection. *Sol. Phys.*, **257**, 217–235.
- DEUBNER, F.-L. (1975) Observations of Low Wavenumber Nonradial Eigenmodes of the Sun. *A&A*, **44**, 371–375.
- DIKPATI, M. & GILMAN, P. A. (2012) Theory of Solar Meridional Circulation at High Latitudes. *ApJ*, **746**, 65.
- DUVALL, JR., T. L. & GIZON, L. (2000) Time-Distance Helioseismology with f-Modes as a Method for Measurement of Near-Surface Flows. *Sol. Phys.*, **192**, 177–191.
- DUVALL, JR., T. L., JEFFERIES, S. M., HARVEY, J. W. & POMERANTZ, M. A. (1993) Time-Distance Helioseismology. *Nature*, **362**, 430–432.
- FISHER, G. H. & WELSCH, B. T. (2008) FLCT: A Fast, Efficient Method for Performing Local Correlation Tracking. In *Subsurface and Atmospheric Influences on Solar Activity*, edited by R. Howe, R. W. Komm, K. S. Balasubramaniam & G. J. D. Petrie, vol. 383 of *Astronomical Society of the Pacific Conference Series*.
- FLECK, B., COUVIDAT, S. & STRAUS, T. (2011) On the Formation Height of the SDO/HMI Fe 6173 Å Doppler Signal. *Sol. Phys.*, **271**, 27–40.
- FLETCHER, L. N., IRWIN, P. G. J., ORTON, G. S., TEANBY, N. A., ACHTERBERG, R. K., BJORAKER, G. L., READ, P. L., SIMON-MILLER, A. A. ET AL. (2008) Temperature and Composition of Saturn’s Polar Hot Spots and Hexagon. *Science*, **319**, 79–81.
- FOURNIER, D., GIZON, L., HOHAGE, T. & BIRCH, A. C. (2014) Generalization of the Noise Model for Time-Distance Helioseismology. *A&A*, **567**, A137.
- FRÖHLICH, C., ROMERO, J., ROTH, H., WEHRLI, C., ANDERSEN, B. N., APPOURCHAUX, T., DOMINGO, V., TELLJOHANN, U. ET AL. (1995) VIRGO: Experiment for Helioseismology and Solar Irradiance Monitoring. *Sol. Phys.*, **162**, 101–128.
- FRUTIGER, C., SOLANKI, S. K., FLIGGE, M. & BRULS, J. H. M. J. (2000) Properties of the Solar Granulation Obtained from the Inversion of Low Spatial Resolution Spectra. *A&A*, **358**, 1109–1121.
- FUHR, J. R., MARTIN, G. A. & WIESE, W. L. (1988) Atomic Transition Probabilities. Iron through Nickel. *Journal of Physical and Chemical Reference Data*, **17**.
- GARCÍA, R. A. (2009) The Sun as a Star: 13 Years of SoHO. In *American Institute of Physics Conference Series*, edited by J. A. Guzik & P. A. Bradley, vol. 1170 of *American Institute of Physics Conference Series*.
- GARCÍA, R. A., TURCK-CHIÈZE, S., JIMÉNEZ-REYES, S. J., BALLOT, J., PALLÉ, P. L., EFF-DARWICH, A., MATHUR, S. & PROVOST, J. (2007) Tracking Solar Gravity Modes: The Dynamics of the Solar Core. *Science*, **316**, 1591–1593.

- GEORGOBIANI, D., ZHAO, J., KOSOVICHEV, A. G., BENSON, D., STEIN, R. F. & NORDLUND, Å. (2007) Local Helioseismology and Correlation Tracking Analysis of Surface Structures in Realistic Simulations of Solar Convection. *ApJ*, **657**, 1157–1161.
- GILMAN, P. A. (1979) Model Calculations concerning Rotation at High Solar Latitudes and the Depth of the Solar Convection Zone. *ApJ*, **231**, 284–292.
- GIZON, L., BALLOT, J., MICHEL, E., STAHN, T., VAUCLAIR, G., BRUNTT, H., QUIRION, P.-O., BENOMAR, O. ET AL. (2013) Seismic Constraints on Rotation of Sun-like Star and Mass of Exoplanet. *Proceedings of the National Academy of Science*, **110**, 13267–13271.
- GIZON, L. & BIRCH, A. C. (2002) Time-Distance Helioseismology: The Forward Problem for Random Distributed Sources. *ApJ*, **571**, 966–986.
- GIZON, L. & BIRCH, A. C. (2004) Time-Distance Helioseismology: Noise Estimation. *ApJ*, **614**, 472–489.
- GIZON, L. & BIRCH, A. C. (2005) Local Helioseismology. *Living Reviews in Solar Physics*, **2**, 6.
- GIZON, L., BIRCH, A. C., BUSH, R. I., DUVAL, JR., T. L., KOSOVICHEV, A. G., SCHERRER, P. H. & ZHAO, J. (2001) Time-Distance Helioseismology and the Solar Orbiter Mission. In *Solar encounter. Proceedings of the First Solar Orbiter Workshop*, edited by B. Battick, H. Sawaya-Lacoste, E. Marsch, V. Martinez Pillet, B. Fleck & R. Marsden., vol. 493 of *ESA Special Publication*.
- GIZON, L., BIRCH, A. C. & SPRUIT, H. C. (2010) Local Helioseismology: Three-Dimensional Imaging of the Solar Interior. *ARA&A*, **48**, 289–338.
- GIZON, L., DUVAL, T. L. & SCHOU, J. (2003) Wave-Like Properties of Solar Supergranulation. *Nature*, **421**, 43–44.
- GIZON, L. & SOLANKI, S. K. (2003) Determining the Inclination of the Rotation Axis of a Sun-like Star. *ApJ*, **589**, 1009–1019.
- GREER, B. J., HINDMAN, B. W., FEATHERSTONE, N. A. & TOOMRE, J. (2015) Helioseismic Imaging of Fast Convective Flows throughout the Near-surface Shear Layer. *ApJ*, **803**, L17.
- HANASOGE, S. M. & BRANICKI, M. (2013) Interpreting Cross-Correlations of One-Bit Filtered Seismic Noise. *Geophysical Journal International*, **195**, 1811–1830.
- HANASOGE, S. M., DUVAL, T. L. & SREENIVASAN, K. R. (2012) Anomalous Weak Solar Convection. *Proceedings of the National Academy of Science*, **109**, 11928–11932.
- HATHAWAY, D. H., UPTON, L. & COLEGROVE, O. (2013) Giant Convection Cells Found on the Sun. *Science*, **342**, 1217–1219.
- HILL, F., BOLDING, J., TONER, C., CORBARD, T., WAMPLER, S., GOODRICH, B., GOODRICH, J., ELIASON, P. ET AL. (2003) The GONG++ Data Processing Pipeline. In *GONG+ 2002. Local and Global Helioseismology: the Present and Future*, edited by H. Sawaya-Lacoste, vol. 517 of *ESA Special Publication*.

- HILL, F., HOWE, R., KOMM, R., GONZALEZ HERNANDEZ, I., TRIPATHY, S. & JAIN, K. (2009) A Helioseismic Comparison of the Solar Minima Preceding Cycles 23 and 24. In *AAS/Solar Physics Division Meeting #40*, vol. 40 of *AAS/Solar Physics Division Meeting*.
- HOWARD, R. & LABONTE, B. J. (1980) The Sun is Observed to be a Torsional Oscillator with a Period of 11 Years. *ApJ*, **239**, L33–L36.
- HOWE, R. (2009) Solar Interior Rotation and its Variation. *Living Reviews in Solar Physics*, **6**, 1.
- HOWE, R., CHRISTENSEN-DALSGAARD, J., HILL, F., KOMM, R., LARSON, T. P., REMPEL, M., SCHOU, J. & THOMPSON, M. J. (2013) The High-latitude Branch of the Solar Torsional Oscillation in the Rising Phase of Cycle 24. *ApJ*, **767**, L20.
- HOWE, R., CHRISTENSEN-DALSGAARD, J., HILL, F., KOMM, R. W., LARSEN, R. M., SCHOU, J., THOMPSON, M. J. & TOOMRE, J. (2000) Dynamic Variations at the Base of the Solar Convection Zone. *Science*, **287**, 2456–2460.
- HUFFMAN, D. A. (1952) A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the I.R.E.*, **40**, 1098–1102.
- JIANG, J., CAMERON, R., SCHMITT, D. & SCHÜSSLER, M. (2009) Countercell Meridional Flow and Latitudinal Distribution of the Solar Polar Magnetic Field. *ApJ*, **693**, L96–L99.
- KATSUKAWA, Y., MASADA, Y., SHIMIZU, T., SAKAI, S. & ICHIMOTO, K. (2010) Pointing Stability of Hinode and Requirements for the next Solar Mission Solar-C. *International Conference on Space Optics*.
- KOSOVICHEV, A. G., CHRISTENSEN-DALSGAARD, J., DAEPPEN, W., DZIEMBOWSKI, W. A., GOUGH, D. O. & THOMPSON, M. J. (1992) Sources of Uncertainty in Direct Seismological Measurements of the Solar Helium Abundance. *MNRAS*, **259**, 536–558.
- KOSOVICHEV, A. G., DUVALL, T. L. & SCHERRER, P. H. (1996) Simulation of Vector-Weighted Binning for the Medium-l Program. *SOI Technical Notes*.
- KOSOVICHEV, A. G., SCHOU, J., SCHERRER, P. H., BOGART, R. S., BUSH, R. I., HOEKSEMA, J. T., ALOISE, J., BACON, L. ET AL. (1997) Structure and Rotation of the Solar Interior: Initial Results from the MDI Medium-L Program. *Sol. Phys.*, **170**, 43–61.
- KUHN, J. R., BUSH, R., EMILIO, M. & SCHOLL, I. F. (2012) The Precise Solar Shape and its Variability. *Science*, **337**, 1638–1640.
- KUMAR, P., SRIVASTAVA, A. K., SOMOV, B. V., MANOHARAN, P. K., ERDÉLYI, R. & UDDIN, W. (2010) Evidence of Solar Flare Triggering Due to Loop-loop Interaction Caused by Footpoint Shear Motion. *ApJ*, **723**, 1651–1664.
- LANGFELLNER, J., GIZON, L. & BIRCH, A. C. (2014) Time-Distance Helioseismology: A new Averaging Scheme for Measuring Flow Vorticity. *A&A*, **570**, A90.

- LANGFELLNER, J., GIZON, L. & BIRCH, A. C. (2015) Spatially Resolved Vertical Vorticity in Solar Supergranulation using Helioseismology and Local Correlation Tracking. *A&A*, accepted.
- LEIBACHER, J. W. & STEIN, R. F. (1971) A New Description of the Solar Five-Minute Oscillation. *Astrophys. Lett.*, **7**, 191–192.
- LEIGHTON, R. B., NOYES, R. W. & SIMON, G. W. (1962) Velocity Fields in the Solar Atmosphere. I. Preliminary Report. *ApJ*, **135**, 474.
- LEKA, K. D., BARNES, G., BIRCH, A. C., GONZALEZ-HERNANDEZ, I., DUNN, T., JAVORNIK, B. & BRAUN, D. C. (2013) Helioseismology of Pre-emerging Active Regions. I. Overview, Data, and Target Selection Criteria. *ApJ*, **762**, 130.
- LINDSEY, C. & BRAUN, D. C. (2000) Seismic Images of the Far Side of the Sun. *Science*, **287**, 1799–1801.
- LISLE, J. & TOOMRE, J. (2004) Cause of Shrinking Sun Effect in Local Correlation Tracking and Impacts on the Mapping of SSW Flows. In *SOHO 14 Helio- and Asteroseismology: Towards a Golden Future*, edited by D. Danesy, vol. 559 of *ESA Special Publication*.
- LÖPTIEN, B., BIRCH, A. C., GIZON, L. & SCHOU, J. (2014a) Image Compression in Local Helioseismology. *A&A*, **571**, A42.
- LÖPTIEN, B., BIRCH, A. C., GIZON, L., SCHOU, J., APPOURCHAUX, T., BLANCO RODRÍGUEZ, J., CALLY, P. S., DOMINGUEZ-TAGLE, C. ET AL. (2014b) Helioseismology with Solar Orbiter. *Space Sci. Rev.*, in press.
- MARSCH, E., ANTONUCCI, E., BOCHSLER, P., BOUGERET, J.-L., HARRISON, R., SCHWENN, R. & VIAL, J.-C. (2000) First Solar Orbiter Assessment Study Report. *ESA-SCI(2000)6*.
- MARSDEN, R. G. & MÜLLER, D. (2011) Solar Orbiter Exploring the Sun-Heliosphere Connection, Solar Orbiter Definition Study Report. *ESA/SRE(2011)14*.
- MARTÍNEZ PILLET, V., DEL TORO INIESTA, J. C., ÁLVAREZ-HERRERO, A., DOMINGO, V., BONET, J. A., GONZÁLEZ FERNÁNDEZ, L., LÓPEZ JIMÉNEZ, A., PASTOR, C. ET AL. (2011) The Imaging Magnetograph eXperiment (IMaX) for the Sunrise Balloon-Borne Solar Observatory. *Sol. Phys.*, **268**, 57–102.
- MÜLLER, D., MARSDEN, R. G., ST. CYR, O. C. & GILBERT, H. R. (2013) Solar Orbiter. Exploring the Sun-Heliosphere Connection. *Sol. Phys.*, **285**, 25–70.
- NAGASHIMA, K., ZHAO, J., KOSOVICHEV, A. G. & SEKII, T. (2011) Detection of Supergranulation Alignment in Polar Regions of the Sun by Helioseismology. *ApJ*, **726**, L17.
- NORDLUND, Å. & STEIN, R. F. (2001) Solar Oscillations and Convection. I. Formalism for Radial Oscillations. *ApJ*, **546**, 576–584.
- NORTON, A. A., GRAHAM, J. P., ULRICH, R. K., SCHOU, J., TOMCZYK, S., LIU, Y., LITES, B. W., LÓPEZ ARISTE, A. ET AL. (2006) Spectral Line Selection for HMI: A Comparison of Fe I 6173 Å and Ni I 6768 Å. *Sol. Phys.*, **239**, 69–91.

- NORTON, A. A. & ULRICH, R. K. (2000) Measuring Magnetic Oscillations in the Solar Photosphere: Coordinated Observations with MDI, ASP and MWO. *Sol. Phys.*, **192**, 403–413.
- NOVEMBER, L. J. & SIMON, G. W. (1988) Precise Proper-Motion Measurement of Solar Granulation. *ApJ*, **333**, 427–442.
- OROZCO SUÁREZ, D. & DEL TORO INIESTA, J. C. (2007) The Usefulness of Analytic Response Functions. *A&A*, **462**, 1137–1145.
- OROZCO SUÁREZ, D., KATSUKAWA, Y. & BELLOT RUBIO, L. R. (2012) The Connection between Internetwork Magnetic Elements and Supergranular Flows. *ApJ*, **758**, L38.
- RIEUTORD, M., ROUDIER, T., ROQUES, S. & DUCOTTET, C. (2007) Tracking Granules on the Sun’s Surface and Reconstructing Velocity Fields. I. The CST Algorithm. *A&A*, **471**, 687–694.
- ROUDIER, T., RIEUTORD, M., PRAT, V., MALHERBE, J. M., RENON, N., FRANK, Z., ŠVANDA, M., BERGER, T. ET AL. (2013) Comparison of Solar Horizontal Velocity Fields from SDO/HMI and Hinode Data. *A&A*, **552**, A113.
- RUZMAIKIN, A. & LINDSEY, C. (2003) Helioseismic Probing of the Solar Dynamo. In *GONG+ 2002. Local and Global Helioseismology: the Present and Future*, edited by H. Sawaya-Lacoste, vol. 517 of *ESA Special Publication*.
- SAMADI, R., NORDLUND, Å., STEIN, R. F., GOUPIL, M. J. & ROXBURGH, I. (2003) Numerical Constraints on the Model of Stochastic Excitation of Solar-Type Oscillations. *A&A*, **403**, 303–312.
- SATO, K. H., GARCIA, R. A., PIRES, S., BALLOT, J., MATHUR, S., MOSSER, B., RODRIGUEZ, E., STARCK, J. L. ET AL. (2010) Inpainting: A Powerful Interpolation Technique for Helio- and Asteroseismic Data. *ArXiv e-prints*.
- SCHERRER, P. H., BOGART, R. S., BUSH, R. I., HOEKSEMA, J. T., KOSOVICHEV, A. G., SCHOU, J., ROSENBERG, W., SPRINGER, L. ET AL. (1995) The Solar Oscillations Investigation - Michelson Doppler Imager. *Sol. Phys.*, **162**, 129–188.
- SCHOU, J. (1999) Migration of Zonal Flows Detected Using Michelson Doppler Imager F-Mode Frequency Splittings. *ApJ*, **523**, L181–L184.
- SCHOU, J. (2003) Wavelike Properties of Solar Supergranulation Detected in Doppler Shift Data. *ApJ*, **596**, L259–L262.
- SCHOU, J., ANTIA, H. M., BASU, S., BOGART, R. S., BUSH, R. I., CHITRE, S. M., CHRISTENSEN-DALSGAARD, J., DI MAURO, M. P. ET AL. (1998) Helioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager. *ApJ*, **505**, 390–417.

- SCHOU, J., SCHERRER, P. H., BUSH, R. I., WACHTER, R., COUVIDAT, S., RABELLO-SOARES, M. C., BOGART, R. S., HOEKSEMA, J. T. ET AL. (2012) Design and Ground Calibration of the Helioseismic and Magnetic Imager (HMI) Instrument on the Solar Dynamics Observatory (SDO). *Sol. Phys.*, **275**, 229–259.
- SEKII, T., APPOURCHAUX, T., FLECK, B. & TURCK-CHIÈZE, S. (2015) Future Mission Concepts for Helioseismology. *Space Sci. Rev.*, in press.
- SNODGRASS, H. B. (1984) Separation of Large-Scale Photospheric Doppler Patterns. *Sol. Phys.*, **94**, 13–31.
- SOLANKI, S. K., BARTHOL, P., DANILOVIC, S., FELLER, A., GANDORFER, A., HIRZBERGER, J., RIETHMÜLLER, T. L., SCHÜSSLER, M. ET AL. (2010) SUNRISE: Instrument, Mission, Data, and First Results. *ApJ*, **723**, L127–L133.
- SOLANKI, S. K., DEL TORO INIESTA, J. C., WOCH, J., GANDORFER, A., HIRZBERGER, J., SCHMIDT, W., APPOURCHAUX, T., ALVAREZ-HERRERO, A. ET AL. (2015) The Polarimetric and Helioseismic Imager for Solar Orbiter: SO/PHI. *ArXiv e-prints*.
- STEIN, R., GEORGIOBIANI, D., TRAMPEDACH, R., LUDWIG, H.-G. & NORDLUND, Å. (2004) Excitation of Radial P-Modes in the Sun and Stars. *Sol. Phys.*, **220**, 229–242.
- STEIN, R. F. & NORDLUND, Å. (2000) Realistic Solar Convection Simulations. *Sol. Phys.*, **192**, 91–108.
- STEIN, R. F. & NORDLUND, Å. (2001) Solar Oscillations and Convection. II. Excitation of Radial Oscillations. *ApJ*, **546**, 585–603.
- THOMPSON, M. J., CHRISTENSEN-DALSGAARD, J., MIESCH, M. S. & TOOMRE, J. (2003) The Internal Rotation of the Sun. *ARA&A*, **41**, 599–643.
- TKACZUK, R., RIEUTORD, M., MEUNIER, N. & ROUDIER, T. (2007) Tracking Granules on the Sun’s Surface and Reconstructing Velocity Fields. II. Error Analysis. *A&A*, **471**, 695–703.
- ULRICH, R. K. (1970) The Five-Minute Oscillations on the Solar Surface. *ApJ*, **162**, 993.
- ŠVANDA, M., ZHAO, J. & KOSOVICHEV, A. G. (2007) Comparison of Large-Scale Flows on the Sun Measured by Time-Distance Helioseismology and Local Correlation Tracking. *Sol. Phys.*, **241**, 27–37.
- VERMA, M., STEFFEN, M. & DENKER, C. (2013) Evaluating Local Correlation Tracking Using CO5BOLD Simulations of Solar Granulation. *A&A*, **555**, A136.
- VÖGLER, A., SHELİYAG, S., SCHÜSSLER, M., CATTANEO, F., EMONET, T. & LINDE, T. (2005) Simulations of Magneto-Convection in the Solar Photosphere. Equations, Methods, and Results of the MURaM code. *A&A*, **429**, 335–351.
- VON DER LÜHE, O. & KENTISCHER, T. J. (2000) High Spatial Resolution Performance of a Triple Fabry-Pérot Filtergraph. *A&AS*, **146**, 499–506.

- WALLACE, G. K. (1992) The JPEG Still Picture Compression Standard. *IEEE Transactions on Consumer Electronics*, **38**, 18–34.
- WELSCH, B. T., FISHER, G. H., ABBETT, W. P. & REGNIER, S. (2004) ILCT: Recovering Photospheric Velocities from Magnetograms by Combining the Induction Equation with Local Correlation Tracking. *ApJ*, **610**, 1148–1156.
- WOCH, J. & GIZON, L. (2007) The Solar Orbiter Mission and its Prospects for Helioseismology. *Astronomische Nachrichten*, **328**, 362.
- ZHAO, J., BOGART, R. S., KOSOVICHEV, A. G., DUVALL, JR., T. L. & HARTLEP, T. (2013) Detection of Equatorward Meridional Flow and Evidence of Double-cell Meridional Circulation inside the Sun. *ApJ*, **774**, L29.
- ZHAO, J., GEORGOBIANI, D., KOSOVICHEV, A. G., BENSON, D., STEIN, R. F. & NORDLUND, Å. (2007) Validation of Time-Distance Helioseismology by Use of Realistic Simulations of Solar Convection. *ApJ*, **659**, 848–857.
- ZHAO, J., NAGASHIMA, K., BOGART, R. S., KOSOVICHEV, A. G. & DUVALL, JR., T. L. (2012) Systematic Center-to-limb Variation in Measured Helioseismic Travel Times and its Effect on Inferences of Solar Interior Meridional Flows. *ApJ*, **749**, L5.







# Publications

## Refereed Publications

- Löptien, B., Birch, A. C., Gizon, L., Schou, J., et al., *Helioseismology with Solar Orbiter*, 2014, *Space Science Reviews*, in press, DOI: 10.1007/s11214-014-0065-3.
- Löptien, B., Birch, A. C., Gizon, L., Schou, J., *Image Compression in Local Helioseismology*, 2014, *Astronomy & Astrophysics*, 571, A42, DOI: 10.1051/0004-6361/201424315.
- Nagashima, K., Löptien, B., Gizon, L., Birch, A. C., Cameron, R., Couvidat, S., Danilovic, S., Fleck B., Stein, R., *Interpreting the Helioseismic and Magnetic Imager (HMI) Multi-Height Velocity Measurements*, 2014, *Solar Physics*, 289, pp. 3457-3481, DOI: 10.1007/s11207-014-0543-5.

## Conference Contributions

- HELAS VI / SOHO 28 / SPACEINN “Helioseismology and Applications”, Göttingen, Germany, Sept. 1-5, 2014  
Talk by B. Löptien: *Helioseismology with Solar Orbiter*
- ISSI workshop “Helioseismology and Dynamics of the Solar Interior”, Bern, Switzerland, Sept. 24-28, 2012  
Invited talk by B. Löptien: *Synthetic helioseismic data for Solar Orbiter*
- 5th Solar Orbiter Workshop, Brugge, Belgium, Sept. 10-13, 2012  
Poster by B. Löptien: *Synthetic SO/PHI data for Helioseismology*

## Submitted Publications

- Löptien, B., Birch, A. C., Duvall Jr., T. L., Gizon, L., Schou, J., *Data Compression for Local Correlation Tracking of Solar Granulation*, 2015, *Astronomy & Astrophysics*, submitted.



# Acknowledgements

First of all, I would like to thank the members of my thesis advisory committee, Aaron C. Birch, Laurent Gizon and Sami K. Solanki. This PhD thesis would not exist without their guidance and support over the last few years.

I thank Laurent Gizon and Manfred Schüssler for refereeing this thesis and I also thank Andreas Tilgner, Ansgar Reiners, Stefan Dreizler, and Jörg Büchner for being in my defense committee.

In addition to my supervisors, many people supported me over the course of this PhD. I am grateful for many fruitful discussions with Jesper Schou. I thank Johann Hirzberger for answering lots of questions about the PHI instrument. I also thank Andreas Lagg and Michiel van Noort for their support regarding the SPINOR code. Many thanks to Robert Stein for providing the STAGGER simulations that I have used extensively during this PhD. I am also grateful to Julián Blanco Rodríguez for providing the SOPHISM code and to Jan Langfellner for providing the tracked and remapped HMI datacubes.

I thank the coordinator of the *International Max Planck Research School for Solar System Science at the University of Göttingen* (IMPRS), Sonja Schuh, especially for helping me with the various bureaucratic steps involved in submitting this thesis.

This work was supported by Deutsche Forschungsgemeinschaft (DFG) through SFB 963/1 "Astrophysical Flow Instabilities and Turbulence" (Project A1). Support was also provided by European Union FP7 projects SPACEINN and SOLARNET. The German Data Center for SDO, funded by the German Aerospace Center (DLR), provided the IT infrastructure for this project. The HMI data used are courtesy of NASA / SDO and the HMI science team.

Last but not least, I would like to thank my family for their moral support.